IEEE Signal Processing
MAGAZINE

SPECTRUM SENSING FOR COGNITIVE RADIO

ACOUSTIC SCENE CLASSIFICATION

DEEP LEARNING IN SPEECH SYNTHESIS

VISUAL DOMAIN ADAPTATION

RESAMPLING METHODS FOR PARTICLE FILTERING

PHASE RETRIEVAL IN OPTICAL IMAGING

PRECIPITATION MONITORING

OPTIMIZATION TOOLS FOR CONFERENCE REVIEW MANAGEMENT

IEEE Signal Processing Society

◆IEEE

CONTENTS

[IEEE **SIGNAL PROCESSING** magazine]

[COVER]

© ISTOCKPHOTO.COM/ELLY99

Printed in the U.S.A.

SUSTAINABLE FORESTRY INITIATIVE
Certified Chain of Custody
At Least 25%Certified Forest Content
www.sfiprogram.org
SFI-01042

SCOPE: *IEEE Signal Processing Magazine* publishes tutorial-style articles on signal processing research and applications, as well as columns and forums on issues of interest. Its coverage ranges from fundamental principles to practical implementation, reflecting the multidimensional facets of interests and concerns of the community. Its mission is to bring up-to-date, emerging and active technical developments, issues, and events to the research, educational, and professional communities. It is also the main Society communication platform addressing important issues concerning all members.

Min Wu
Editor-in-Chief
minwu@umd.edu

[from the EDITOR]

Impact Beyond Numbers

When you receive this issue of *IEEE Signal Processing Magazine* (*SPM*), the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) will be taking place in beautiful Brisbane, Australia. *SPM's* Editorial Board will meet in person during ICASSP. This is a valuable opportunity for the Editorial Board to reflect on the progress made so far, the plans being carried out, and to brainstorm ideas to bring the magazine to the next level.

Ten new Editorial Board members started their term this year: Sven Lončarić (University of Zagreb, Croatia), Brian Lovell (University of Queensland, Australia), Yi Ma (ShanghaiTech University, China), Henrique (Rico) Malvar (Microsoft Research), Athina Petropulu (Rutgers University), Peter Ramadge (Princeton University), Shigeki Sagayama (Meiji University and emeritus University of Tokyo, Japan), Shihab Shamma (University of Maryland), Gregory Wornell (Massachusetts Institute of Technology), and Dapeng Wu (University of Florida). Together with the continuing Editorial Board members, these colleagues have brought to our magazine a tremendous amount of collective knowledge and experiences. Knowing the many commitments that they already have, I greatly appreciate their willingness to serve on *SPM's* Editorial Board.

I would also like to welcome Dr. Andres Kwasinski, who was a devoted area editor for columns and forum for the past three years, as our area editor for social media and outreach. This newly created area editor position will help explore new types of content and provide effective outreach to members and readers.

The magazine has been a premier platform for researchers to contribute tutorial surveys and overviews on the latest advances in signal processing. This issue of *SPM* includes three clusters of feature articles centered on learning and classification, new advances in signal processing theories and methods, and interesting new signal processing applications. It is due to the tireless efforts of Prof. Marc Moonen, past area editor for feature articles, and Prof. Abdelhak Zoubir, *SPM's* past editor-in-chief, that we are able to bring this diverse set of articles to you in one issue. Prof. Shuguang (Robert) Cui, *SPM's* new area editor for feature articles, also contributed to assembling this issue. My sincere thanks to all of their efforts!

It is common today to characterize the impact of articles using citation statistics. Here, beyond numbers, I would like to share a personal experience of publishing with *SPM* that may shine some light toward the impact on authors and readers. My first article with *SPM* was in response to the call for papers to the special issue on digital rights management (DRM) more than ten years ago. I was working with several colleagues on tracing the leak of multimedia documents by embedding specially designed signals in image and video so that each copy is uniquely labeled. The guest editors reminded us of *SPM's* tutorial article style, which was in place to ensure that articles were to be understood by a broad audience.

One of the guidelines that I still remember today is the number of equations—no more than three—which sounded impossible at first: after all, we were planning to synthesize the work from a series of research papers by several representative groups, and the number of equations in each of these papers was in the double digits! This seemingly stringent constraint pushed us to think hard on how to present the ideas in accessible terms, with the minimum number of equations. For example, to explain the essential idea of a complex code construction from a seminal theoretical work, we developed a toy example and created step-by-step illustrations. This process of publishing a tutorial article with *SPM* helped me develop a deeper understanding toward the research problems and obtain valuable insights that inspired later research.

The article was published in the March 2004 issue of *SPM* as part of a timely and balanced article collection on DRM with beautiful artistic designs. The IEEE *Xplore* online library was in its infancy then. So I mailed hard copies of the issue to several researchers overseas, including one to Prof. Yanda Li, who led the signal and information processing program at my college alma mater, Tsinghua University, in China. Later that year, I received a phone call from a college friend with whom I hadn't been in contact for many years. As it turned out, this friend faced an antipiracy challenge when developing digital technologies for China's broadcasting industry, but few researchers in China at the time had worked on this problem. When he came to consult Prof. Li, my article in that special issue provided a starting point for discussion. The world is so small! Indeed, beyond citation numbers, *SPM* has served as a vehicle to connect researchers across mountains and oceans, and bring together signal processing professionals in academia and industry.

[SP]

Alex Acero
2014–2015 SPS President
a.acero@ieee.org

president's **MESSAGE**

The IEEE Gives Our Society the "Thumbs Up"

In January 2014, shortly after starting my term as president of the IEEE Signal Processing Society (SPS), I learned it was time to prepare the IEEE Technical Activities Board (TAB) five-year Society review report. Each member of the SPS's Executive Committee was responsible for at least one section of the report; however, I was responsible for quite a few. After several months of work, we ended up with a 100-page-long document. Fortunately, our great staff led us through the process, and it ended up being much less daunting than it initially looked. We delivered our report to the IEEE Society Review Committee during the June 2014 IEEE TAB meeting held in New Jersey. We recently received the committee's feedback and, in a nutshell, we passed with flying colors!

The Review Committee praised the Society's large portfolio of publications (the SPS is the fourth-largest Society in the IEEE in terms of members but the second in terms of the number of journal and magazine articles published), number of conferences [having created the China Summit and International Conference on Signal and Information Processing (ChinaSIP) and the Global Conference on Signal and Information Processing (Global-SIP) during the last five years], and our sound finances. They were impressed by the new initiatives under our membership board: the Chapter of the Year Award, the Chapter certification process, SigView, SigPort, and the seasonal schools on emerging topics. Finally, they also appreciated the creation of special interest groups and our nascent effort to encourage volunteers to author Wikipedia pages on signal processing topics. These efforts must have paid off—Society membership has increased to over 17,000 members from 13,800 in 2009 (few Societies have

> **WE RECENTLY RECEIVED THE COMMITTEE'S FEEDBACK AND, IN A NUTSHELL, WE PASSED WITH FLYING COLORS!**

grown that much in the same period), including the highest percentage growth in student membership (+17.7% since January 2013). The Review Committee said they'd be passing along our Society's best practices to other Societies.

Roughly half of the Society's 17,000 members are from industry, but the percentage of industry members on our Executive Committee and Board of Governors is significantly lower. Accordingly, the Review Committee encouraged us to look for ways to increase industry participation in the governing boards of the Society. They also encouraged us to investigate ways to involve industry members more, possibly though local workshops in the Chapters, while suggesting that more surveys could be offered to our industry members.

The IEEE reviews each of the 45 Societies and councils every five years. While I won't be involved five years from now when it's our turn to do it again, I think the Society review is a great mechanism for control and feedback. In fact, a decade ago, we decided to set up a similar mechanism, patterned after the five-year Society review, to review our technical committees.

Our Society's Executive Committee is composed of a president; president-elect; and vice presidents for publications, conferences, technical directions, and membership. We have monthly conference calls, meet at ICASSP (the International Conference on Acoustics, Speech, and Signal Processing) and yet again in the fall [a time that coincides with GlobalSIP or ICIP (the International Conference on Image Processing)], and answer numerous e-mails. We all have day jobs in academia or industry, so our volunteer time tends to be spent running the day-to-day tasks of the Society. Although preparing for the five-year Society review was a lot of work, it provided us with a great opportunity to step back and get a little perspective, and we intend to follow up on all of the review committee's recommendations. Preparing for the review really taught me what a great Society we have and its potential for growth., and I'd love to hear from all of you how we can make things even better.

Acero

[SP]

Intrinsically Hopeful

In November 2014, Prof. Thomas Kailath from Stanford University was presented with the National Medal of Science by U.S. President Barack Obama in Washington, D.C., during a ceremony honoring ten of the top American scientists and engineers. The medal was established by U.S. President Dwight Eisenhower in 1959, and the first medal was awarded by U.S. President John F. Kennedy in 1963.

Prof. Kailath is a Life Fellow of the IEEE and has been an IEEE Signal Processing Society member for more than 40 years. Throughout his career, he made significant contributions to signal processing. In the 1960s, he was mostly interested in signal detection before turning onto signal estimation in the 1970s. The ESPRIT algorithm is a well-known outcome of this line of research. In the 1980s, Prof. Kailath then focused on various aspects of array processing as well as the design of very-large-scale integration architectures for signal processing applications. His research team has, for instance, developed spatial multiplexing in multiple-input, multiple-output antenna systems, which is now used in Wi-Fi. In the 1990s, signal processing ideas were instrumental in his work on optical microlithography, when his team broke what was believed to be the 100-nm barrier in semiconductor manufacturing by Gordon Moore and
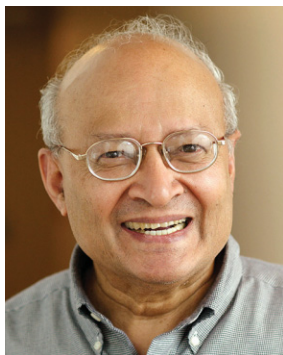
L.A. CICERO/STANFORD NEWS SERVICE

several others. Some of these contributions are still standard industry practice at the present time. His technical achievements have been acknowledged over the years with top awards from both the IEEE Signal Processing Society and the IEEE Information Theory Society.

During the National Medal of Science ceremony, Prof. Kailath was recognized for "transformative contributions to the fields of information and system science, for distinctive and sustained mentoring of young scholars, and for translation of scientific ideas into entrepreneurial ventures that have had a significant impact on industry." A short remark by President Obama gave an even more

personal flavor to the award ceremony. "As Thomas Kailath, one of our honorees today, says, 'Scientists are intrinsically hopeful and believe in grand answers, and that if we work hard enough we can find some of them in our lifetime.' And that's a good phrase: *intrinsically hopeful*. I'm intrinsically hopeful, I am [laughter]. That's who I am. That's who we are as a people, as Americans, as a nation." According to Prof. Kailath, the quotation originates from an offline discussion with a member of the staff who was asking about failures when he was talking about his contributions. He replied that he could not really recollect any major failure and then elaborated further saying the words quoted by the president or something close to it. Possibly, it resonated with the president because *hope* had been a major theme of his first election campaign.

[SP]

THOMAS KAILATH

Current Job: Hitachi America Professor of Engineering Emeritus, Stanford University, United States

Birthplace: Pune, India

Education: B.E. degree (1956) from the University of Pune, India; S.M. degree (1959) and Sc.D. degree (1961) from the Massachusetts Institute of Technology, United States

First Job: Counting seeds in a seed-packing facility

Major Awards: U.S. National Medal of Science; IEEE Medal of Honor; IEEE Jack S. Kilby Signal Processing Medal; U.S. National Academy of Engineering and National Academy of Sciences; and many more.

Learn More on the Web:
- Stanford Univ. [Online]. Available: http://web.stanford.edu/~tkailath/
- Mini-Documentary by National Medals Foundation. [Online]. Available: http://www.youtube.com/watch?v=58n2ONrcCRw
- T. S. Perry, "Medal of Honor: Thomas Kailath," *IEEE Spect.* [Online]. Available: http://spectrum.ieee.org/computing/networks/medal-of-honor-thomas-kailath. doi: 10.1109/MSPEC.2007.352532

3rd IEEE Global Conference on Signal & Infomation Processing
Orlando, Florida, USA December 14-16 2015

IEEE GlobalSIP'15–Call for Papers   (http://2015.ieeeglobalsip.org/)

General Chairs: Jose Moura and Dapeng Oliver Wu
Technical Program Chairs: Mihaela van der Schaar, Xiaodong Wang, and Hsiao-Chun Wu

The IEEE Global Conference on Signal and Information Processing (GlobalSIP) is a recently launched flagship conference of the IEEE Signal Processing Society. GlobalSIP' 15 will be held in Orlando, Florida, USA, December 14–16, 2015. The conference will focus broadly on signal and information processing with an emphasis on up-and-coming signal processing themes. The conference will feature world-class speakers, tutorials, exhibits, and technical sessions consisting of poster or oral presentations. GlobalSIP' 15 technical program will be comprised of a main program (General Symposium) and several co-located symposia on special topics. Technical paper submissions are solicited in the interest topics, which may include, but are not limited to:

- Signal processing in communications and networks, including green communication and signal processing in optical communication
- Image and video processing
- Selective topics in speech and language processing
- Signal processing in security applications
- Signal processing in energy and power systems

- Signal processing in genomics and bioengineering (physiological, pharmacological and behavioral)
- Signal processing for social media networks
- Neural signal processing
- Seismic signal processing
- Hardware and real-time implementations
- Other novel and significant applications of selected areas of signal processing

Symposia:
- General Symposium
- Symposium on Signal Processing on Graphics Processing Units and Multicores
- Symposium on Signal Processing in Mobile Multimedia Communication Systems
- Symposium on 3GPP EVS and Beyond
- Symposium on Signal and Information Processing for Optimizing Future Energy Systems
- Symposium on Signal Processing Challenges in Human Brain Connectomics
- Symposium on Real-Time Signal Processing for Low-Cost and Low-Power Smart Devices
- Symposium on Signal Processing for Optical Wireless Communications
- Symposium on Signal and Information Processing for Software-Defined Ecosystems, and Green Computing
- Symposium on Signal Processing Applications in Smart Buildings

Submission of Papers: Prospective authors are invited to submit full-length papers, with up to four pages for technical content including figures and possible references, and with one additional optional 5th page containing only references. Manuscripts should be original (not submitted/published anywhere else) and written in accordance with the standard IEEE double-column paper template. All paper submissions should be carried out through EDAS system (http://edas.info). A selection of best papers and best student papers will be made by the GlobalSIP 2015 best paper award committee upon recommendations from Technical Committees.

Timeline for paper submission:

May 15, 2015:	Paper submission deadline
June 30, 2015:	Review results announced
September 5, 2015:	Camera-ready papers due

Notice: The IEEE Signal Processing Society enforces a "no-show" policy. Any accepted paper included in the final program is expected to have at least one author or qualified proxy attend and present the paper at the conference. Authors of the accepted papers included in the final program who do not attend the conference will be subscribed to a "No-Show List", compiled by the Society. The "no-show" papers will not be published by IEEE on IEEE *Xplore* or other public access forums, but these papers will be distributed as part of the on-site electronic proceedings and the copyright of these papers will belong to the IEEE.

[reader's **CHOICE**]

Top Downloads in IEEE *Xplore*

The "Reader's Choice" column in *IEEE Signal Processing Magazine* contains a list of articles published by the IEEE Signal Processing Society (SPS) that ranked among the top 100 most downloaded IEEE *Xplore* articles. This issue is based on download data through December 2014. The table below contains the citation information for each article and the rank obtained in IEEE *Xplore*. The highest rank obtained by an article in this time frame is indicated in bold. Your suggestions and comments are welcome and should be sent to Associate Editor Michael Gormish (gormish@ieee.org).

TITLE, AUTHOR, PUBLICATION YEAR IEEE SPS PUBLICATIONS	ABSTRACT	RANK IN IEEE TOP 100						*N* TIMES IN TOP 100 (SINCE JAN 2011)
		DEC 2014	NOV 2014	OCT 2014	SEP 2014	AUG 2014	JUL 2014	
GENERALIZED NEW MERSENNE NUMBER TRANSFORMS Boussakta, S.; Hamood, M.T.; Rutter, N. *IEEE Transactions on Signal Processing* vol. 60, no. 5, 2012, pp. 2640–2647	Two new number theoretic transforms named as odd and odd-squared new Mersenne number transforms are introduced. An example is given which shows their suitability for the calculation of different types of convolutions and other algorithms.	**14**						1
LESSONS FOR RADAR Vespe, M.; Jones, G.; Baker, C.J. *IEEE Signal Processing Magazine* vol. 26, no. 1, 2009, pp. 65–75	A range of strategies employed by bats is considered for possible exploitation in the radar systems of tomorrow. Focus is given to the functions necessary for autonomous navigation.	**33**						1
AN OVERVIEW OF MASSIVE MIMO: BENEFITS AND CHALLENGES Lu, L.; Li, G.Y.; Swindlehurst, A.L.; Ashikhmin, A.; Zhang, R. *IEEE Journal on Selected Topics in Signal Processing* vol. 8, no. 5, 2014, pp. 742–758	Equipping cellular base stations with a very large number of antennas potentially allows for orders of magnitude improvement in spectral and energy efficiency. This paper presents an extensive overview and analysis of massive MIMO systems.	**37**			66			3
IMAGE QUALITY ASSESSMENT: FROM ERROR VISIBILITY TO STRUCTURAL SIMILARITY Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. *IEEE Transactions on Image Processing* vol. 13, no. 4, 2004, pp. 600–612	This paper introduces a framework for quality assessment based on the degradation of structural information. Within this framework a structure similarity index is developed and evaluated. MATLAB code is available.	59	29	33	45	25	**17**	27
K-SVD: AN ALGORITHM FOR DESIGNING OVERCOMPLETE DICTIONARIES FOR SPARSE REPRESENTATION Aharon, M.; Elad, M.; Bruckstein, A. *IEEE Transactions on Signal Processing* vol. 54, no. 11, 2006, pp.4311–4322	K-SVD is an iterative method that alternates between sparse coding of the examples based on the current dictionary and a process of updating the dictionary atoms to better fit the data in a computationally efficient manner.	61	39	**38**	53	48	60	8
WEIGHTED GUIDED IMAGE FILTERING Li, Z.; Zheng, J.; Zhu, Z.; Yao, W.; Wu, S. *IEEE Transactions on Image Processing* vol. 24, no. 1, 2015, pp. 120–129	The weighted guided image filter incorporates an edge-aware weighting into existing guided image filter to address the problem of halo artifacts. The filter is applied to detail enhancement, haze removal and image fusion.	**64**						1

TITLE, AUTHOR, PUBLICATION YEAR IEEE SPS PUBLICATIONS	ABSTRACT	RANK IN IEEE TOP 100						N TIMES IN TOP 100 (SINCE JAN 2011)
		DEC 2014	NOV 2014	OCT 2014	SEP 2014	AUG 2014	JUL 2014	
MODELING AND OPTIMIZATION FOR BIG DATA ANALYTICS: (STATISTICAL) LEARNING TOOLS FOR OUR ERA OF DATA DELUGE Slavakis, K.; Giannakis, G.B.; Mateos, G. *IEEE Signal Processing Magazine* vol. 31, no. 5, 2014, pp. 18–31	This article offers scalable architectures and optimization algorithms for decentralized and online learning problems for SP-relevant tasks such as PCA, dictionary learning, compressive sampling, and subspace clustering.	78	62	**52**	**52**			4
IMAGE SUPER-RESOLUTION VIA SPARSE REPRESENTATION Yang, J.; Wright, J.; Huang, T.S.; Ma, Y. *IEEE Transactions on Image Processing* vol. 19, no. 11, 2010, pp. 2861–2873	This paper presents an approach to single-image super-resolution based upon sparse signal representation of low and high-resolution patches.	91	83	89			**45**	16
MODULATION FORMATS AND WAVEFORMS FOR 5G NETWORKS: WHO WILL BE THE HEIR OF OFDM? Banelli, P.; Buzzi, S.; Colavolpe, G.; Modenini, A.; Rusek, F.; Ugolini, A. *IEEE Signal Processing Magazine* vol. 31, no. 6, 2014, pp. 80–93	This article provides a review of some modulation formats suited for 5G enriched by a comparative analysis of their performance in a cellular environment, and by a discussion on their interactions with specific 5G ingredients.	95	**80**	91				3
IMAGE QUALITY ASSESSMENT FOR FAKE BIOMETRIC DETECTION: APPLICATION TO IRIS, FINGERPRINT, AND FACE RECOGNITION Galbally, J.; Marcel, S.; Fierrez, J. *IEEE Transactions on Image Processing* vol. 23, no. 2, 2014, pp. 710–724	This paper uses 25 general image quality features extracted from the authentication image to distinguish between legitimate and imposter samples for fingerprint, iris, and 2D face biometrics.	97		69	48	**15**	40	8
SCALING UP MIMO: OPPORTUNITIES AND CHALLENGES WITH VERY LARGE ARRAYS Rusek, F.; Persson, D.; Lau, B.K.; Larsson, E.G.; Marzetta, T.L; Edfors, O; Tufvesson, F. *IEEE Signal Processing Magazine* vol. 30, no. 1, 2013, pp. 40–60	The more antennas the transmitter/ receiver is equipped with and the more degrees of freedom that the propagation channel can provide, the better the performance in terms of data rate or link reliability. This article quantifies the reliability and achievable rates.		**67**	79				17
NEW CHALLENGES FOR IMAGE PROCESSING RESEARCH Pappas, T.N. *IEEE Transactions on Image Processing* vol. 20, no. 12, 2011, p. 3321	The editor-in-chief of *IEEE Transactions on Image Processing* addresses the direction of the journal and image processing.		87		80	**23**	27	5
CONVEX OPTIMIZATION FOR BIG DATA: SCALABLE, RANDOMIZED, AND PARALLEL ALGORITHMS FOR BIG DATA ANALYTICS Cevher, V.; Becker, S.; Schmidt, M. *IEEE Signal Processing Magazine* vol. 31, no. 5, 2014, pp. 32–43	This article reviews recent advances in convex optimization algorithms for big data, which aim to reduce the computa- tional, storage, and communications bottlenecks.			**55**	75	92		3
PRIVACY PRESERVING DATA SHARING WITH ANONYMOUS ID ASSIGNMENT Dunning, L.A.; Kresman, R. *IEEE Transactions on Information Forensics and Security* vol. 8, no. 2, 2013, pp. 402–413	This paper offers an algorithm and analyzes multiple algorithms to assign ID numbers ranging from 1 to N to N parties without using a trusted central authority and is still resistant to collusion among other members.			59		**28**		3
A TUTORIAL ON PARTICLE FILTERS FOR ONLINE NONLINEAR/NON-GAUSSIAN BAYESIAN TRACKING Arulampalam, M.S.; Maskell, S.; Gordon, N.; Clapp, T. *IEEE Transactions on Signal Processing* vol. 50, no. 2, 2002, pp. 174–188	This paper reviews optimal and suboptimal Bayesian algorithms for nonlinear/ non-Gaussian tracking problems, with a focus on particle filters. Variants of the particle filter are introduced within a framework of the sequential importance sampling (SIS) algorithm and compared with the standard EKF.			86	**77**		80	42
SUPER-RESOLUTION IMAGE RECON- STRUCTION: A TECHNICAL OVERVIEW Park, S.C.; Park, M.K.; Kang, M.G. *IEEE Signal Processing Magazine* vol. 20, no. 3, 2003, pp. 21–36	This article introduces the concept of super-resolution (SR) algorithms and presents a technical review of various existing SR methodologies and models the low-resolution image acquisition process.			93		**56**		19

[reader's **CHOICE**]

TITLE, AUTHOR, PUBLICATION YEAR IEEE SPS PUBLICATIONS	ABSTRACT	RANK IN IEEE TOP 100						_N_ TIMES IN TOP 100 (SINCE JAN 2011)
		DEC 2014	NOV 2014	OCT 2014	SEP 2014	AUG 2014	JUL 2014	
RECENT DEVELOPMENTS IN THE SPARSE FOURIER TRANSFORM: A COMPRESSED FOURIER TRANSFORM FOR BIG DATA Gilbert, C.; Indyk, P.; Iwen, M.; Schmidt, L. _IEEE Signal Processing Magazine_ vol. 31, no. 5, 2014, pp. 91–100	This article surveys recent developments of the sparse Fourier transform, which addresses big data issues by computing a compressed Fourier transform using only a subset of the input data, in time smaller than the data set size.			94	85	90		3
LOW COMPLEXITY EQUALIZATION FOR DOUBLY SELECTIVE CHANNELS MODELED BY A BASIS EXPANSION Hrycak, T.; Das, S.; Matz, G.; Feichtinger, H.G. _IEEE Transactions on Signal Processing_ vol. 58, no. 11, 2010, pp. 5706–5719	The equalizer computes a regularized solution of a linear system involving the channel matrix, which utilizes the product-convolution structure without ever explicitly creating the channel matrix. The proposed equalizer achieves BERs comparable to those of MMSE equalization, and outperforms low-complexity equalizers.				24			1
BIG DATA ANALYSIS WITH SIGNAL PROCESSING ON GRAPHS: REPRESENTATION AND PROCESSING OF MASSIVE DATA SETS WITH IRREGULAR STRUCTURE Sandryhaila, A.; Moura, J.M.F. _IEEE Signal Processing Magazine_ vol. 31, no. 5, 2014, pp. 80–90	Fundamental concepts of discrete signal processing on graphs including graph signals and graph filters, graph Fourier transform, graph frequency, and spectrum ordering are reviewed and compared with counterparts from classical signal processing. Product graphs are used as a graph model to extend to large data.				87			1
ORDINAL FEATURE SELECTION FOR IRIS AND PALMPRINT RECOGNITION Sun, Z.; Wang, L.; Tan, T. _IEEE Transactions on Image Processing_ vol. 23, no. 9, 2014, pp. 3922–3934	Feature selection is designed to achieve an accurate and sparse representation of ordinal measures. Formulation as a linear programming problem obtains efficient quality results on the CASIA and PolyU databases.				62			1
VECTOR-VALUED IMAGE PROCESSING BY PARALLEL LEVEL SETS Ehrhardt, M.J.; Arridge, S.R. _IEEE Transactions on Image Processing_ vol. 23, no. 1, 2014, pp 9–18	Considers the components of an image as a vector. By minimizing large angles, parallel level sets are obtained and used for demosaicking.						35	5
GRADIENT HISTOGRAM ESTIMATION AND PRESERVATION FOR TEXTURE ENHANCED IMAGE DENOISING Zuo, W.; Zhang, L.; Song, C.; Zhang, D.; Gao, H. _IEEE Transactions on Image Processing_ vol. 23, no. 6, 2014, pp. 2459–2472	This paper avoids the smoothing associated with many denoising algorithms by preserving the histogram of gradients in an image. Region based variants handle different textures.						89	1

[SP]

John Edwards

Signal Processing Enhances Environmental Sensing

Sensors and other data sources, combined with sophisticated signal processing techniques, promise to help scientists better observe and analyze various types of environmental data.

Biologist Nathan Merchant, for example, has created a method for tracking ships and monitoring underwater noise levels in protected marine mammal habitats. Merchant, senior scientist for underwater noise at the U.K. Centre for Environment, Fisheries, and Aquaculture Science (CEFAS), developed the system with coresearchers Enrico Pirotta, Tim Barton, and Paul Thompson, of The Institute of Biological and Environmental Sciences at Scotland's University of Aberdeen.

"Underwater noise levels have risen significantly over time in step with human activity," Merchant says. These changes in the acoustic environment affect marine mammals, which rely on sound as their primary sensory mode. "The disturbance caused by man-made noise can disrupt crucial activities, such as hunting for food, affecting the animals' health."

To help understand the impact noise might exert on dolphins and their population levels, the researchers conducted a study on Moray Firth, Scotland's largest inlet (Figure 1). Moray Firth is home to a population of bottlenose dolphins as well as numerous types of seals, porpoises, and whales. The protected habitat also hosts construction yards that supply Scotland's rapidly expanding offshore wind farm industry. Projected increases in wind farm construction are expected to drive more shipping through the habitat—something many scientists believe could eventually negatively impact resident marine mammals.

"Various types of ships emit noise at different levels and frequencies, therefore it is vital to know which kinds of vessels are crossing the habitats and migration routes of marine mammals," Merchant says. Merchant and his fellow researchers recently monitored underwater noise levels using hydrophone sensors (underwater microphones), ship-tracking data, and shore-based time-lapse photography. The techniques created a ship-noise assessment toolkit.

"In this project, we used signal processing techniques to integrate several different data sources into one package: time-lapse video, underwater sound recordings, and ship-tracking data," Merchant says. Using signal processing techniques in combination with video editing software, the researchers combined video, audio, and spatial data to produce a synchronized audiovisual representation of the soundscape, including shipping activity and weather conditions across the marine mammal habitat.

The main challenge the researchers faced was processing each data source in a way that would supply a common time resolution for the audiovisual output. "We used geolocation tools to map the ship tracking data through time, which involved temporal and spatial interpolation of the raw data, which has a fairly coarse, ~10 minute time resolution," Merchant explains. "The sound recordings were averaged at intervals that corresponded to the time resolution of the video and spatial data, and an adaptive thresholding algorithm was also developed which detected when a ship was passing."

The approach, Merchant says, allowed the researchers to link diverse data sources and gain insights into the habitat's sonic environment in a way that would not have been possible by interpreting each

data source individually. "The integrative approach produced a ship noise assessment analysis that was much more than the sum of its parts," Merchant adds.

One of the main difficulties in studying underwater sound is that long-term recordings generate vast amounts of acoustic data—several terabytes in the case of this project. "Consequently, we have had to develop high-performance computing techniques for processing big data, which involved using parallel processing across many cores of a large server or cluster," Merchant says. "Now that we have these in place, we can process large datasets rapidly, but getting there was quite a challenge."

Merchant notes that the biggest challenge he currently faces is developing models to predict how sound will spread through an underwater area, which would enable the researchers to produce maps of sound levels in a particular habitat. "This [capability] can be used in environmental impact assessment of noisy activities, like offshore wind farm construction, because it shows us over how big an area marine mammals could potentially be disturbed," Merchant says. "We are currently refining and testing scripts using data from several field studies."

According to Merchant, there is an almost endless number of coastal areas where shipping interacts with nearby habitats. "Not only... marine mammals but also fishes and invertebrates, which we are aware are also sensitive to noise," he says. "The techniques that have been developed in this project can very much be applied to assess what kind of ships are making noise, what kind of noise level is generated, and how they are concentrated spatially, as well as how all of this interacts with the habitats."

Merchant also believes that the new techniques will spin off applications extending far beyond marine habitat

[special **REPORTS**] continued

[FIG1] Researchers deploy a hydrophone sound sensor in Scotland's Moray Firth as a dolphin surfaces nearby. (Photo courtesy of U.K. CEFAS.)

protection. "There will potentially be a lot of people interested in being able to, for example, detect when ships are passing along a corridor that's on a tracking system," he adds. "We are aware that there might be some interest in doing that kind of thing [from] Coast Guard and military types."

FOLLOWING URBAN VIBRATIONS

Urban traffic—including cars, trucks, trains, and planes—generates both acoustic and seismic noise. While most people can easily detect vehicle noise, seismic vibrations are usually not perceptible to humans. Nevertheless, a pair of researchers at the Scripps Institution of Oceanography at the University of California at San Diego believe that seismic "noise" could soon become a useful data source for next-generation traffic information systems.

While the detection of naturally occurring seismic vibrations has long been useful to scientists searching for subsurface features like earthquake faults and petroleum resources, the various types of vibrations generated by traffic flows have never been explored in any real depth, says Nima Riahi, a Scripps postdoctoral fellow working alongside Peter Gerstoft, a Scripps geophysicist. The pair believes that a future urban seismic network could tap into

vehicle-generated vibrations to monitor the flow of human transport across a specific area.

Last year, energy company Signal Hill Petroleum of Signal Hill, California, gave the researchers access to a large vibration data set covering the area under the city of Long Beach, California. "We seized the opportunity," explains Riahi. The data set—mapped by a 5,300-geophone network—was as part of a hydrocarbon industry survey covering an area of more than 7×10 km (Figure 2). Geophones are devices commonly used by private, government and academic researchers to record energy waves reflected by subsurface geology, typically as a way of mapping out geologic structures or tracking earthquakes.

"By analyzing vibrations from geophones spaced approximately 100 m (300 ft) apart, we were able to examine activity in Long Beach with a resolution below a typical city block," Riahi adds. He notes that the spatiotemporal structure of the man-made seismic noise intensity revealed individual train activity along the area's Blue Line Metro railway line, allowed the counting of departing and landing aircraft at Long Beach Airport (as well as estimating their motion) and gave

clues about traffic movement along Interstate 405, a major southern California freeway. More advanced analysis techniques and algorithms promise to reveal many other types of manmade signals within the ground, Riahi says, potentially leading to the monitoring of activities beyond traffic flow.

"The findings indicate that human seismic noise might serve as a rich data source for the observation of cities," Riahi says. "The approach could also be used for urban area characterization, allowing various types and schedules of activities to be visualized, making it possible to vibrationally identify specific industrial, residential or office zones."

Riahi describes the research accomplished so far as "simple and straightforward" signal processing. "We tried to keep it simple at first. It is essentially calculating the power of the vibration as a function of time." A custom-design spatiotemporal filter was also used to remove vibrations that failed to match a pattern indicating a type of ongoing movement, such as a train traveling down a track.

The researchers are only interested in examining various types of continuous vibrations, which exist in many different variations. "There are a lot of things happening: day/night variations, trucks passing, which might be different than when a car passes," Riahi says. "We want to see if there are similarities between different things; can we group things together, like in cluster analysis?"

Freeway traffic proved to be more difficult to discern and analyze than train or airport movements. "The 405 is challenging because it is a ten-lane highway, two directions," Riahi says. "We had about 13 sensors per kilometer of highway—that is really a low spatial sampling." Yet, although they were restricted to only a limited number of sensors, the researchers were still able to detect individual trucks moving along the roadbed at night. "We know that, because there is a continuous motion detected from one sensor to the next at about 55 miles per hour going through the entire stretch of the 405 section we were looking at," Riahi explains.

Finding seismic needles in a geological haystack required Riahi and Gerstoft to

[FIG2] An aerial view of Long Beach, California, showing a portion of the 5,300-geophone network. Interstate 405 runs through the photo's center; a Long Beach Airport runway is on the right. (Photo courtesy of Scripps Institution of Oceanography at the University of California at San Diego.)

consider a wide range of approaches. "Clustering techniques are an interesting path to pursue when you are just trying to look for structure in the data," Riahi remarks. The team investigated the potential of various clustering algorithms. "One of them, obviously, is *K*-means, which is a popular but nonoptimal clustering algorithm," Riahi says. "There are also algorithms based on sparse coding and sparse reconstruction, where you are saying, 'I have some signal, I think it is composed of a few elemental components and I am trying to find out which one it is.'"

According to Riahi, the spatiotemporal filter required significant creativity. "I had to custom-write that filter because I was not aware of anything that would work for our data," Riahi says. "There were other options; I tried image processing filters, for instance, but the ones that I came across and tried out did not work so well."

Riahi says the study showed that anthropogenic seismic power—a relatively simple attribute—when analyzed with a dense grid of urban seismic sensors, can measure a wide range of human activities. "The human imprint on the seismic wave field provides a rich, but so far underappreciated, data source to observe cities," Riahi remarks.

MEASURING SEA LEVELS FROM SPACE

A new way of measuring sea levels, developed by researchers at Sweden's Chalmers University of Technology, promises to generate faster and more accurate readings. Measuring sea level is an important part of climate research, since a rising mean sea level is a key indicator of climate change.

Johan Löfgren and Rüdiger Haas, research scientists at Chalmers' Department of Earth and Space Sciences, have created a Global Navigation Satellite System (GNSS) tide gauge, an instrument that measures sea level by using radio signals from satellite navigation systems. "We want to be able to make detailed measurements of sea level so that we can understand how coastal societies will be affected in the future," Löfgren says.

The GNSS tide gauge uses radio signals from Earth-orbiting satellites within satellite navigation systems like global positioning system (GPS) and Glonass (Russia's equivalent of GPS). Two antennas measure signals directly from the satellites and signals reflected off the sea surface (Figure 3). By analyzing these signals together, the sea level and its variation can be measured up to 20 times per second.

[FIG3] When a satellite passes overhead, the GNSS tide gauge uses signals from the satellite and signals reflected off the sea surface to measure the current sea level. (Photo courtesy of Onsala Space Observatory/J. Löfgren.)

The GNSS tide gauge has an advantage over previous technologies in that it can measure changes in both land and sea simultaneously in the same location. Therefore, both long-term and short-term land movements can be taken into

Acoustic Scene Classification

[Classifying environments from the sounds they produce]

[
Daniele Barchiesi,
Dimitrios Giannoulis,
Dan Stowell, and
Mark D. Plumbley
]

n this article, we present an account of the state of the art in acoustic scene classification (ASC), the task of classifying environments from the sounds they produce. Starting from a historical review of previous research in this area, we define a general framework for ASC and present different implementations of its components. We then describe a range of different algorithms submitted for a data challenge that was held to provide a general and fair benchmark for ASC techniques. The data set recorded for this purpose is presented along with the performance metrics that are used to evaluate the algorithms and statistical significance tests to compare the submitted methods.

We use a baseline method that employs Mel-frequency cepstral coefficients (MFCCs), Gaussian mixture models (GMMs), and a maximum likelihood criterion as a benchmark and only find sufficient evidence to conclude that three algorithms significantly outperform it. We also evaluate the human classification accuracy in performing a similar classification task. The best-performing algorithm achieves a mean accuracy that matches the median accuracy obtained by humans, and common pairs of classes are misclassified by both computers and humans. However, all acoustic scenes are correctly classified by at least some individuals, while there are scenes that are misclassified by all algorithms.

INTRODUCTION

Enabling devices to make sense of their environment through the analysis of sounds is the main objective of research in machine listening, a broad investigation area related to computational auditory scene analysis (CASA) [51]. Machine-listening systems perform

© ISTOCKPHOTO.COM/AGSANDREW

analogous processing tasks to the human auditory system and are part of a wider research theme linking fields such as machine learning, robotics, and artificial intelligence.

ASC refers to the task of associating a semantic label to an audio stream that identifies the environment in which it has been produced. Throughout the literature on ASC, a distinction is made between psychoacoustic/psychological studies aimed at understanding the human cognitive processes that enable our understanding of acoustic scenes [35] and computational algorithms that attempt to automatically perform this task using signal processing and machine-learning methods. The perceptual studies have also been referred to as *soundscape cognition* [15] by defining soundscapes as the auditory equivalent of landscapes [43]. In contrast, the computational research has also been called *computational auditory scene recognition* [38]. This is a task related to the area of CASA [51] and is particularly applied to the study of environmental sounds [18]. It is worth noting that, although many ASC studies are inspired by biological processes, ASC algorithms do not necessarily employ frameworks developed within CASA and the two research fields do not completely overlap.

Work in ASC has evolved in parallel with several related research problems. For example, methods for the classification of noise sources have been employed for noise-monitoring systems [22] or to enhance the performance of speech processing algorithms [17]. Algorithms for sound source recognition [13] attempt to identify the sources of acoustic events in a recording and are closely related to event detection and classification techniques. The latter methods are aimed at identifying and labeling temporal regions containing single events of a specific class and have been employed, e.g., in surveillance systems [40], elderly assistance [26], and speech analysis through the segmentation of acoustic scenes [29]. Furthermore, algorithms for the semantic analysis of audio streams that also rely on the recognition or clustering of sound events have been used for personal archiving [19] and audio segmentation [33] and retrieval [53].

The distinction between event detection and ASC can sometimes appear blurred, e.g., when considering systems for multimedia indexing and retrieval [9], where the identification of events, such as the sound produced by a baseball player batting in a run, also characterizes the general environment (in this case, the environment of a baseball game). On the other hand, ASC can be employed to enhance the performance of sound event detection [28] by providing prior information about the probability of certain events. To limit the scope of this article, we will only detail systems aimed at modeling complex physical environments containing multiple events.

Applications that can specifically benefit from ASC include the design of context-aware services [45], intelligent wearable devices [52], robotics navigation systems [11], and audio archive management [32]. Concrete examples of possible future technologies that could be enabled by ASC include smartphones that continuously sense their surroundings, switching to silent mode every time a person enters a concert hall; assistive technologies such as hearing aids or robotic wheelchairs that adjust their functioning based on the recognition of indoor or outdoor environments; or sound archives that automatically assign metadata to audio files. Moreover, classification could be performed as a preprocessing step to inform algorithms developed for other applications, such as source separation of speech signals from different types of background noise. Although this article details methods for the analysis of audio signals, it is worth mentioning that, to address the aforementioned problems, acoustic data can be combined with other sources of information such as geolocation, acceleration sensors, collaborative tagging, and filtering.

From a purely scientific point of view, ASC represents an interesting problem that both humans and machines are only able to solve to a certain extent. From the outset, semantic labeling of an acoustic scene or soundscape is a task open to different interpretations, as there is not a comprehensive taxonomy encompassing all the possible categories of environments. Researchers generally define a set of categories, record samples from these environments,

and treat ASC as a supervised classification problem within a closed universe of possible classes. Furthermore, even within pre-defined categories, the set of acoustic events or qualities character-izing a certain environment is generally unbounded, making it difficult to derive rules that unambiguously map acoustic events or features to scenes.

BACKGROUND: A HISTORY OF ASC

The first method appearing in the literature to specifically address the ASC problem was proposed by Sawhney and Maes in a 1997 technical report from the Massachusetts Institute of Technology (MIT) Media Lab [42]. The authors recorded a data set from a set of classes including people, voices, the subway, and traffic. They extracted several features from the audio data using tools borrowed from speech analysis and auditory research, employing recurrent neural networks and a k-nearest neighbor criterion to model the mapping between features and categories, and obtaining an overall classification accuracy of 68%. One year later, researchers from the same institution recorded a continuous audio stream by wearing a microphone while making a few bicycle trips to a supermarket and then automatically segmented the audio into different scenes (such as a home, a street, and a supermarket) [12]. For the classifi-cation, they fitted the empirical distribution of features extracted from the audio stream to hidden Markov models (HMMs).

Meanwhile, research in experimental psychology was focused on understanding the perceptual processes driving the human ability to categorize and recognize sounds and soundscapes. Bal-las found that the speed and accuracy in the recognition of sound events is related to the acoustic nature of the stimuli, how often they occur, and whether they can be associated with a physical cause or a sound stereotype [4]. Peltonen et al. observed that the human recognition of soundscapes is guided by the identification

of typical sound events, such as human voices or car engine noises, and measured an overall 70% accuracy in the human abil-ity to discern among 25 acoustic scenes [37]. Dubois et al. investi-gated how individuals define their own taxonomy of semantic categories when this is not given a priori by the experimenter [15]. Finally, Tardieu et al. tested both the emergence of semantic classes and the recognition of acoustic scenes within the context of rail stations [47]. They reported that sound sources, human activities, and room effects such as reverberation are the ele-ments driving the formation of soundscape classes and the cues employed for recognition when the categories are fixed a priori.

Influenced by the psychoacoustic/psychological literature that emphasized both local and global characteristics for the recogni-tion of soundscapes, some of the computational systems that built on the early works by researchers at MIT [42], [12] focused on modeling the temporal evolution of audio features. Eronen et al. employed MFCCs to describe the local spectral envelope of audio signals and GMMs to describe their statistical distribution [21]. Next, they trained HMMs to account for the temporal evolu-tion of the GMMs using a discriminative algorithm that exploited knowledge about the categories of training signals. Eronen et al. further developed this work by considering a larger group of fea-tures and adding a feature transform step to the classification algorithm, obtaining an overall 58% accuracy in the classification of 18 different acoustic scenes [20].

In the algorithms mentioned so far, each signal belonging to a training set of recordings is generally divided into frames of fixed duration, and a transform is applied to each frame to obtain a sequence of feature vectors. The feature vectors derived from each acoustic scene are then employed to train a statistical model that summarizes the properties of a whole soundscape or of mul-tiple soundscapes belonging to the same category. Finally, a

ACRONYM	AUTHORS	TITLE
RNH	G. ROMA, W. NOGUEIRA, AND P. HERRERA	RECURRENCE QUANTIFICATION ANALYSIS FEATURES FOR AUDITORY SCENE CLASSIFICATION
RG	A. RAKOTOMAMONJY AND G. GASSO	HISTOGRAM OF GRADIENTS OF TIME–FREQUENCY REPRESENTATIONS FOR AUDIO SCENE CLASSIFICATION
GSR	J.T. GEIGER, B. SCHULLER, AND G. RIGOLL	RECOGNIZING ACOUSTIC SCENES WITH LARGE-SCALE AUDIO FEATURE EXTRACTION AND SVM
CHR	M. CHUM, A. HABSHUSH, A. RAHMAN, AND C. SANG	IEEE AASP SCENE CLASSIFICATION CHALLENGE USING HMMs AND FRAME-BASED CLASSIFICATION
NHL	J. NAM, Z. HYUNG, AND K. LEE	ASC USING SPARSE FEATURE LEARNING AND SELECTIVE MAX-POOLING BY EVENT DETECTION
NR	W. NOGUEIRA, G. ROMA, AND P. HERRERA	SOUND SCENE IDENTIFICATION BASED ON MFCC, BINAURAL FEATURES AND A SUPPORT VECTOR MACHINE (SVM) CLASSIFIER
PE	K. PATIL AND M. ELHILALI	MULTIRESOLUTION AUDITORY REPRESENTATIONS FOR SCENE CLASSIFICATION
KH	J. KRIJNDERS AND G.A.T. HOLT	A TONE-FIT FEATURE REPRESENTATION FOR SCENE CLASSIFICATION
ELF	B. ELIZALDE H. LEI, G. FRIEDLAND, AND N. PETERS	AN i-VECTOR-BASED APPROACH FOR AUDIO SCENE DETECTION
LTT*	D. LI, J. TAM, AND D. TOUB	AUDITORY SCENE CLASSIFICATION USING MACHINE-LEARNING TECHNIQUES
OE	E. OLIVETTI	THE WONDERS OF THE NORMALIZED COMPRESSION DISSIMILARITY REPRESENTATION

[TABLE 1] THE LIST OF ALGORITHMS SUBMITTED FOR THE DCASE CHALLENGE ON ASC.

* The original LTT submission achieved low accuracy due to a bug in a MATLAB toolbox—here we are presenting the results obtained with the correct implementation.

decision criterion is defined to assign unlabeled recordings to the category that best matches the distribution of their features.

FEATURES

Several categories of audio features have been employed in ASC systems. Here, we present a list and provide their rationale in the context of audio analysis for classification. [Here and throughout the article, the notation [1, XXX] (see Table 1) is used to cite the extended abstracts submitted for the detection and classification of acoustic scenes and events (DCASE) challenge described in "Challenge on Detection and Classification of Acoustic Scenes and Events" section.]

1) *Low-level time-based and frequency-based audio descriptors*: Several ASC systems [1, GSR] [20], [34] employ features that can be easily computed from either the signal in the time domain or its Fourier transform. These include (among others) the zero crossing rate, which measures the average rate of sign changes within a signal and is related to the main frequency of a monophonic sound; the spectral centroid, which measures the center of mass of the spectrum and is related to the perception of brightness [25]; and the spectral roll-off that identifies a frequency above which the magnitude of the spectrum falls below a set threshold.

2) *Frequency-band energy features* (*energy/frequency*): This class of features used by various ASC systems [1, NR CHR GSR], [20] is computed by integrating the magnitude spectrum or the power spectrum over specified frequency bands. The resulting coefficients measure the amount of energy present within different subbands and can also be expressed as a ratio between the subband energy and the total energy to encode the most prominent frequency regions in the signal.

3) *Auditory filter banks*: A further development of energy/frequency features consists of analyzing audio frames through filter banks that mimic the response of the human auditory system. Sawhney and Maes used Gammatone filters for this purpose [42]. Clarkson et al. instead computed Mel-scaled filter bank coefficients (MFCs) [12], whereas Patil and Elahili [1, PE] employed a so-called auditory spectrogram.

4) *Cepstral features*: MFCCs are an example of cepstral features and are perhaps the most popular features used in ASC. They are obtained by computing the discrete DCT of the logarithm of MFCs. The word *cepstral* is an anagram of the word *spectral* and indicates that this class of features is computed by applying a Fourier-related transform to the spectrum of a signal. Cepstral features capture the spectral envelope of a sound and, thus, summarize their coarse spectral content.

5) *Spatial features*: If the soundscape has been recorded using multiple microphones, features can be extracted from the different channels to capture the properties of the acoustic scene. In the case of a stereo recording, popular features include the interaural time difference (ITD), which measures the relative delay occurring between the left and right channels when recording a sound source, and the interaural level difference (ILD), which measures the amplitude variation between channels. Both ITD and ILD are linked to the

position of a sound source in the stereo field. Nogueira et al. included spatial features in their ASC system [1, NR].

6) *Voicing features*: Whenever the signal is thought to contain harmonic components, a fundamental frequency f_0 or a set of fundamental frequencies can be estimated, and groups of features can be defined to measure the properties of these estimates. In the case of ASC, harmonic components might correspond to specific events occurring within the audio scene, and their identification can help discriminate between different scenes. Geiger et al. employed voicing features related to the fundamental frequency of each frame in their system [1, GSR]. The method proposed by Krijnders and Holt [1, KH] is based on extracting tone-fit features, a sequence of voicing features derived from a perceptually motivated representation of the audio signals. First, a so-called cochleogram is computed to provide a time–frequency representation of the acoustic scenes inspired by the properties of the human cochlea. Then, the tonalness of each time–frequency region is evaluated to identify tonal events in the acoustic scenes, resulting in tone-fit feature vectors.

7) *Linear predictive coefficients* (*LPCs*): This class of features has been employed in the analysis of speech signals that are modeled as autoregressive processes. In an autoregressive model, samples of a signal s at a given time instant t are expressed as linear combinations of samples at L previous time instants

$$s(t) = \sum_{l=1}^{L} \alpha_l s(t-l) + \varepsilon(t), \qquad (1)$$

where the combination coefficients $\{\alpha_l\}_{l=1}^{L}$ determine the model parameters and ϵ is a residual term. There is a mapping between the value of LPCs and the spectral envelope of the modeled; therefore, α_l encodes information regarding the general spectral characteristics of a sound. Eronen et al. employed LPC features in their proposed method [20].

8) *Parametric approximation features*: Autoregressive models are a special case of approximation models where a signal s is expressed as a linear combination of J basis functions from the set $\{\varphi_j\}_{j=1}^{J}$

$$s(t) = \sum_{j=1}^{J} \alpha_j \phi_j(t) + \epsilon(t). \qquad (2)$$

Whenever the basis functions ϕ_j are parameterized by a set of parameters γ_j, features can be defined according to the functions that contribute to the approximation of the signal. For example, Chu et al. decompose audio scenes using the Gabor transform, which is a representation where each basis function is parameterized by its frequency f, its time scale u, its time shift τ, and its frequency phase θ so that $\gamma_j = \{f_j, u_j, \tau_j, \theta_j\}$ [10]. The set of indexes identifying nonzero coefficients $j^* = \{j : \alpha_j \neq 0\}$ corresponds to a set of active parameters γ_{j^*} contributing to the approximation of the signal and encodes events in an audio scene that occur at specific time–frequency locations. Patil and Elahili also extract parametric features derived from the two-dimensional (2-D) convolution between the auditory spectrogram and 2-D Gabor filters [1, PE].

9) *Unsupervised learning features*: The model (2) assumes that a set of basis functions is defined a priori to analyze a signal. Alternatively, bases can be learned from the data or from other features already extracted in an unsupervised way. Nam et al. employed a sparse restricted Boltzmann machine (SRBM) to adaptively learn features from the MFCCs of the training data [1, NHL]. An SRBM is a neural network that has been shown to learn basis functions from input images, which resemble the properties of representations built by the visual receptors in the human brain. In the context of ASC, an SRBM adaptively encodes basic properties of the spectrum of the training signals and returns a sequence of features learned from the MFCCs along with an activation function that is used to determine time segments containing significant acoustic events.

10) *Matrix factorization methods*: The goal of matrix factorization for audio applications is to describe the spectrogram of an acoustic signal as a linear combination of elementary functions that capture typical or salient spectral elements and are, therefore, a class of unsupervised learning features. The main intuition that justifies using matrix factorization for classification is that the signature of events that are important in the recognition of an acoustic scene should be encoded in the elementary functions, leading to discriminative learning. Cauchi employed nonnegative matrix factorization (NMF) [8], and Benetos et al. used probabilistic latent component analysis in their proposed algorithms [6]. Note that a matrix factorization also outputs a set of activation functions that encode the contribution of elementary functions in time, hence modeling the properties of a whole soundscape. Therefore, this class of techniques can be considered to jointly estimate local and global parameters.

11) *Image processing features*: Rakotomamonjy and Gasso designed an algorithm for ASC whose feature extraction function comprises the following operations [1, RG]. First, the audio signals corresponding to each training scene are processed using a constant-Q transform, which returns frequency representations with logarithmically spaced frequency bands. Then, 512×512-pixel grayscale images are obtained from the constant-Q representations by interpolating neighboring time–frequency bins. Finally, the features are extracted from the images by computing the matrix of local gradient histograms. This is obtained by dividing the images into local patches, defining a set of spatial orientation directions, and counting the occurrence of edges exhibiting each orientation. Note that, in this case, the vectors of features are not independently extracted from frames but from time–frequency tiles of the constant-Q transform.

12) *Event detection and acoustic unit descriptors*: Heittola et al. proposed a system for ASC that classifies soundscapes based on a histogram of events detected in a signal [27]. During the training phase, the occurrence of manually annotated events (such as a honking car horn, applause, or a basketball bouncing) is used to derive models for each scene category. In the test phase, HMMs are employed to identify events within an unlabeled recording and to define a histogram that is compared to the ones derived from the

training data. This system represents an alternative to the common framework that includes features, statistical learning, and a decision criterion in that it essentially performs event detection and ASC at the same time. However, for the purpose of this tutorial, the acoustic events can be thought of as high-level features whose statistical properties are described by histograms.

A similar strategy is employed by Chaudhuri et al. to learn acoustic unit descriptors (AUDs) and classify YouTube multimedia data [9]. AUDs are modeled using HMMs and used to transcribe an audio recording into a sequence of events. The transcriptions are assumed to be generated by N-gram language models whose parameters are trained on different soundscape categories. The transcriptions of unlabeled recordings during the test phase are, thus, classified following a maximum likelihood criterion.

FEATURE PROCESSING
The features described so far can be further processed to derive new quantities that are used either in place or as an addition to the original features.

FEATURE TRANSFORMS
This class of methods is used to enhance the discriminative capability of features by processing them through linear or nonlinear transforms. Principal component analysis (PCA) is perhaps the most commonly cited example of feature transforms. It learns a set of orthonormal bases that minimize the Euclidean error resulting from projecting the features onto subspaces spanned by subsets of the basis set (the principal components) and, hence, identifies the directions of maximum variance in the data set. Because of this property, PCA and the more general independent component analysis (ICA) have been employed as dimensionality reduction techniques to project high-dimensional features onto lower-dimensional subspaces while retaining the maximum possible amount of variance [1, PE] [20], [34]. Nogueira et al., on the other hand, evaluate a Fisher score to measure how features belonging to the same class are clustered near each other and far from features belonging to different classes [1, NR]. A high Fisher score implies that features extracted from different classes are likely to be separable, and it is used to select optimal subsets of features.

TIME DERIVATIVES
For all of the quantities computed on local frames, discrete time derivatives between consecutive frames can be included as additional features that identify the time evolution of the properties of an audio scene.

STATISTICAL MODELS
Once the features are extracted from the audio frames, the next stage of an ASC system generally consists of learning statistical models of the distribution of the features. Statistical models are parametric mathematical models used to summarize the properties of individual audio scenes or whole soundscape categories

MFCCs, GMMs, AND A MAXIMUM LIKELIHOOD CRITERION

MFCCs

MCCs have been widely used as a feature for audio analysis. Let $s_n \in \mathbb{R}^D$ be a signal frame and $|\hat{s}_n|$ the absolute value of its Fourier transform. The coefficients corresponding to linearly spaced frequency bins are mapped onto R Mel frequency bands to approximate the human perception of pitches (which can be approximately described as *logarithmic*, meaning that we are capable of a much better resolution at low frequencies than at high frequencies), resulting in $L \leq D$ coefficients. The magnitude of the Mel coefficients is converted to a logarithmic scale and the resulting vector is processed using a discrete cosine transform (DCT). Finally, the $K \leq R$ first coefficients are selected and constitute the vector of features $x_n = \mathcal{T}(s_n)$. This last step essentially measures the frequency content of the log-magnitude of the spectrum of a signal and, therefore, captures general properties of the spectral envelope. For example, periodic sounds that exhibit spectral peaks at multiples of a fundamental frequency are highly correlated with one or several cosine bases, encoding this information in the value of the corresponding MFCC coefficients. The set of parameters $\theta = \{D, R, K\}$ includes frames, dimension, the number of Mel bands, and the number of DCT coefficients that need to be defined when computing the MFCCs. These parameters determine the dimensionality reduction introduced by the features extraction operator, and their choice is governed by the trade-off between generalization and discrimination mentioned in the section "A General Framework for ASC."

STATISTICAL NORMALIZATION

To classify features extracted from signals belonging to different categories, it is important to evaluate the relative differences between the values of feature vectors belonging to different classes rather than differences between different coefficients within feature vectors extracted from the same signal. For this reason, during the training phase of the ASC classification algorithm, statistical normalization is performed as a standard feature processing aimed at avoiding offsets or scaling variations of any of the coefficients within feature vectors. This is accomplished by subtracting the global mean (computed from features extracted from the whole data set) from each vector $x_{n,m}$ and by dividing each coefficient by their global standard deviation. After the feature vectors have been normalized, the average and standard deviation of the coefficients $x_{n,m,k}$ are 0 and 1, respectively.

GMMs

GMMs are used to infer global statistical properties of the features from local features vectors, which are interpreted as realizations of a generative stochastic process. Let $\mathcal{N}(\mu, \Sigma)$ be a multivariate normal distribution with mean $\mu \in \mathbb{R}^K$ and

covariance matrix $\Sigma \in \mathbb{R}^{K \times K}$, and recall that the notation $x_{n,\Lambda q}$ identifies features vectors extracted from training signals that belong to the qth category. Then, every such vector is modeled as generated by the following distribution:

$$x_{n,\Lambda q} \sim \prod_{i=1}^{I} w_i \mathcal{N}(\mu_i, \Sigma_i), \tag{S1}$$

where I is a fixed number of components and w_i is a latent variable expressing the probability that a particular observation is generated from the ith component.

The operator $\mathcal{S}$ takes the collection of features $x_{n,\Lambda q}$ and learns a global model for the qth class $\mathcal{M}_q = \{w_i, \mu_i, \Sigma_i\}_{i=1}^{I}$ by estimating the parameters of the Gaussian mixture distribution in (S1), which can be accomplished through an expectation-maximization (EM) algorithm [7]. The only parameter to be set in this case is the number of Gaussian components I, which rules a tradeoff between model accuracy and overfitting. Indeed, $\mathcal{S}_I$ must include a sufficient number of components to account for the fact that different events within a soundscape generate sounds with different spectral properties. However, as the number of components becomes too large, the model tends to fit spurious random variations in the training data, hindering the generalization capabilities of the algorithm when confronted with an unlabeled sound.

MAXIMUM LIKELIHOOD CRITERION

Once the GMMs' M_q have been inferred from the training data, features can be extracted from an unlabeled sound by applying the operator $\mathcal{T}$. The new sequence of features $x_{n,\text{new}}$ is statistically normalized using the same mean and standard deviation values obtained from the training signals, and a likelihood measure $\mathcal{G}$ is employed to evaluate which class is statistically most likely to generate the observed features, hence determining the sound classification. A set of coefficients g_q is computed by evaluating the log-likelihood of the observed data given the model

$$g_q = p(x_{n,\text{new}} | \mathcal{M}_q) \propto \sum_{i=1}^{I} w_i (x_{n,\text{new}} - \mu_i)^T \Sigma_i (x_{n,\text{new}} - \mu_i), \tag{S2}$$

and a category is picked based on the most likely model $c^*_{\text{new}} = \arg\min_q g_q$.

Note that the baseline system described here is an example of a bag-of-frames technique where the ordering of the sequence of features is irrelevant. Any random permutation of the sequences $x_{n,\Lambda q}$ does not affect the computation of the GMM parameters and, thus, the classification of unlabeled signals.

from the feature vectors. They can be divided into generative or discriminative methods.

When working with generative models, feature vectors are interpreted as being generated from one of a set of underlying statistical distributions. During the training stage, the parameters

of the distributions are optimized based on the statistics of the training data. In the test phase, a decision criterion is defined to determine the most likely model that generated a particular observed example. A simple implementation of this principle is to compute the basic statistical properties of the distribution of

feature vectors belonging to different categories (such as their mean values), hence obtaining one class centroid for each category. The same statistic can be computed for each unlabeled sample that is assumed to be generated according to the distribution with the closest centroid and is assigned to the corresponding category.

When using a discriminative classifier, the features derived from an unlabeled sample are not interpreted as being generated by a class-specific distribution but are assumed to occupy a class-specific region in the feature space. One of the most popular discriminative classifiers for ASC is the SVM. The model output from an SVM determines a set of hyperplanes that optimally separate features associated to different classes in the training set (according to a maximum-margin criterion). An SVM can only discriminate between two classes. However, when the classification problem includes more than two categories (as is the case in the ASC task presented in this article), multiple SVMs can be combined to determine a decision criterion that allows for discrimination between Q classes. In the one-versus-all approach, Q SVMs are trained to discriminate between data belonging to one class and data from the remaining $Q - 1$ classes. Instead, in the one-versus-one approach, $Q(Q-1)/2$ SVMs are trained to classify between all possible class combinations. In both cases, the decision criterion estimates the class from an unlabeled sample by evaluating the distance between the data and the separating hyperplanes learned by the SVMs.

Discriminative models can be combined with generative ones. For example, one might use the parameters of generative models learned from training data to define a feature space and then employ an SVM to learn separating hyperplanes. In other words, discriminative classifiers can be used to derive classification criteria from either the feature vectors or the parameters of their statistical models. In the former case, the overall classification of an acoustic scene must be decided from the classification of individual data frames using, e.g., a majority vote.

Different statistical models have been used for computational ASC, and the following list highlights their categories.

1) *Descriptive statistics*: Several techniques for ASC [1, KH GSR RNH] employ descriptive statistics. This class of methods is used to quantify various aspects of statistical distributions, including moments (such as mean, variance, skewness, and kurtosis of a distribution), quantiles, and percentiles.

2) *GMMs*: Other methods for ASC [11], [2] employ GMMs, which are generative methods where feature vectors are interpreted as being generated by a multimodal distribution expressed as a sum of Gaussian distributions. GMMs are further detailed in "MFCCs, GMMs, and a Maximum Likelihood Criterion" where we will present a baseline ASC system used for a benchmark.

3) *HMMs*: This class of models is used in several ASC systems [12], [20] to account for the temporal unfolding of events within complex soundscapes. Suppose, for example, that an acoustic scene recorded in an underground train includes an alert sound preceding the sound of the doors closing and the noise of the electric motor moving the carriage to the next station. The features extracted from these three distinct

sounds could be modeled using Gaussian densities with different parameters, and the order in which the events normally occur would be encoded in an HMM transition matrix. This contains the transition probability between different states at successive times, which is the probability of each sound occurring after the other.

A transition matrix that correctly models the unfolding of events in an underground train would contain large diagonal elements indicating the probability of sounds persisting in time, significant probabilities connecting events that occur after each other (an alert sound followed by the sound of the doors closing and then the sound of motors), and negligible probabilities connecting sounds that occur in the wrong order (for example, the doors closing before the alert sound).

4) *Recurrence quantification analysis*: Roma et al. employ recurrence quantification analysis (RQA) to model the temporal unfolding of acoustic events [1, RNH]. This technique is used to learn a set of parameters that have been developed to study dynamical systems in the context of chaos theory and are derived from so-called recurrence plots, which capture periodicities in a time series. In the context of ASC, the RQA parameters include: recurrence measuring the degree of self-similarity of features within an audio scene; determinism, which is correlated to sounds periodicities; and laminarity, which captures sounds containing stationary segments. The outputs of the statistical learning function are a set of parameters that model each acoustic scene in the training set. This collection of parameters is then fed to an SVM to define the decision boundaries between classes that are used to classify unlabeled signals.

5) *i-vector*: The system proposed by Elizalde et al. [1, ELF] is based on the computation of the i-vector [14]. This is a technique originally developed in the speech processing community to address a speaker verification problem, and it is based on modeling a sequence of features using GMMs. In the context of ASC, the i-vector is specifically derived as a function of the parameters of the GMMs learned from MFCCs. It leads to a low-dimensional representation summarizing the properties of an acoustic scene and is input into a generative probabilistic linear discriminant analysis (pLDA) [30].

DECISION CRITERIA

Decision criteria are functions used to determine the category of an unlabeled sample from its feature vectors and from the statistical model learned from the set of training samples. Decision criteria are generally dependent on the type of statistical learning methods used. The the following details how different models are associated to the respective criteria.

1) *One versus one and one versus all*: These decision criteria are associated to the output of a multiclass SVM and are used to map the position of a features vector to a class, as already described in the section "Statistical Models."

2) *Majority vote*: This criterion is used whenever a global classification must be estimated from decisions about single audio frames. Usually, an audio scene is classified according

to the most common category assigned to its frames. Alternatively, a weighted majority vote can be employed to vary the importance of different frames. Patil and Elahili, for example, assign larger weights to audio frames containing more energy [1, PE].

3) *Nearest neighbor*: According to this criterion, a feature vector is assigned to the class associated to the closest vector from the training set (according to a metric, often the Euclidean distance). A generalization of the nearest neighbor is the k-nearest neighbor criterion, whereby the k closest vectors are considered and a category is determined according to the most common classification.

4) *Maximum likelihood*: This criterion is associated with generative models, whereby feature vectors are assigned to the category whose model is most likely to have generated the observed data according to a likelihood probability.

5) *Maximum a posteriori* (*MAP*): An alternative to maximum likelihood classification is the MAP criterion, which includes information regarding the marginal likelihood of any given class. For instance, suppose a global positioning system in a mobile device indicates that, in the current geographic area, some environments are more likely to be encountered than others. This information could be included in an ASC algorithm through an MAP criterion.

META-ALGORITHMS

In the context of supervised classification, meta-algorithms are machine-learning techniques designed to reduce the classification error by running multiple instances of a classifier in parallel, each of which uses different parameters or different training data. The results of each classifier are then combined into a global decision.

DECISION TREES AND TREE BAGGERS

A decision tree is a set of rules derived from the analysis of features extracted from training signals. It is an alternative to generative and discriminative models because it instead optimizes a set of if/else conditions about the values of features that leads to a classification output. Li et al. employed a tree-bagger classifier, which is a set of multiple decision trees [1, LTT]. A tree bagger is an example of a classification meta-algorithm that computes

multiple so-called weak learners (classifiers whose accuracy is only assumed to be better than chance) from randomly sampled copies of the training data following a process called *bootstrapping*. In the method proposed by Li et al., the ensemble of weak learners are then combined to determine a category for each frame and, in the test phase, an overall category is assigned to each acoustic scene based on a majority vote.

NORMALIZED COMPRESSION DISSIMILARITY AND RANDOM FOREST

Olivetti adopts a system for ASC that departs from the techniques described throughout this article in favor of a method based on audio compression and random forest [1, OE]. Motivated by the theory of Kolmogorov complexity, which measures the shortest binary program that outputs a signal and that is approximated using compression algorithms, he defines a normalized compression distance between two audio scenes. This is a function of the size in bits of the files obtained by compressing the acoustic scenes using any suitable audio coder. From the set of pairwise distances, a classification is obtained using a random forest, which is a meta-algorithm based on decision trees.

MAJORITY VOTE AND BOOSTING

The components of a classification algorithm can themselves be thought of as parameters subject to optimization. Thus, a further class of meta-algorithms deals with selecting from or combining multiple classifiers to improve the classification accuracy. Perhaps the simplest implementation of this general idea is to run several classification algorithms in parallel on each test sample and determine the optimal category by a majority vote, an approach that will be also used in the section "Evaluation of Algorithms for ASC." Other more sophisticated methods include boosting techniques [44], where the overall classification criterion is a function of linear combinations involving a set of weak learners.

A GENERAL FRAMEWORK FOR ASC

Now that we have seen the range of machine-learning and signal processing techniques used in the context of ASC, let us define a framework that allows us to distill a few key operators and components. Computational algorithms for ASC are designed to solve a
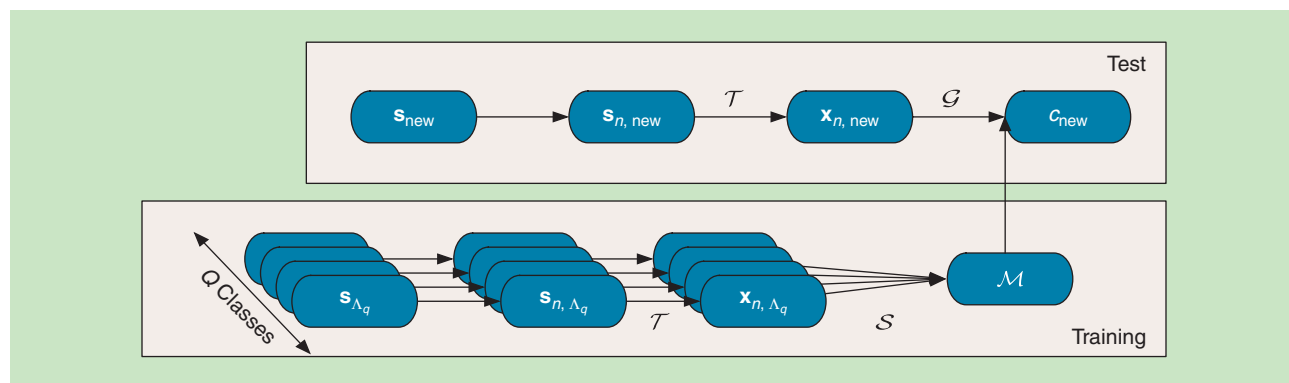
[FIG1] A supervised classification framework for ASC.

supervised classification problem where a set of M training recordings $\{s_m\}_{m=1}^{M}$ is provided and associated with corresponding labels $\{c_m\}_m^M = 1$ that indicate the category to which each soundscape belongs. Let $\{\gamma_q\}_{q=1}^{Q}$ be a set of labels indicating the members of a universe of Q possible categories. Each label c_m can assume one of the values in this set, and we define a set, $\Lambda_q = \{m : c_m = \gamma_q\}$, that identifies the signals belonging to the qth class. The system learns statistical models from the different classes during an off-line training phase and uses them to classify unlabeled recordings, s_{new}, in the test phase.

First, each of the training signals is divided into short frames. Let D be the length of each frame, and $s_{n,m} \in \mathbb{R}^D$ indicates the nth frame of the mth signal. Typically, D is chosen so that the duration of the frames is about 50 ms depending on the signal's sampling rate.

The frames in the time domain are not directly employed for classification but rather are used to extract a sequence of features through a transform $\mathcal{T} : \mathcal{T}(s_{n,m}) = x_{n,m}$, where $x_{n,m} \in \mathbb{R}^K$ indicates a vector of features of dimension K. Often, $K \ll D$, meaning that $\mathcal{T}$ causes a dimensionality reduction. This is aimed at obtaining a coarser representation of the training data where members of the same class result in similar features (yielding generalization) and members of different classes can be distinguished from each other (allowing discrimination). Some systems further manipulate the features using feature transforms, such as in the method proposed by Eronen et al. [20]. For clarity of notation, we will omit this additional feature processing step from the description of the ASC framework, considering any manipulation of the features to be included in the operator $\mathcal{T}$.

The individual features obtained from time-localized frames cannot summarize the properties of soundscapes that are constituted by a number of different events occurring at different times. For this reason, sequences of features extracted from signals belonging to a given category are used to learn statistical models of that category, abstracting the classes from their empirical realizations. Let x_{n,Λ_q} indicate the features extracted from the signals belonging to the qth category. The function $\mathcal{S} : \mathcal{S}(\{x_{n,\Lambda_q}\}) = \mathcal{M}$ learns the parameters of a statistical model $\mathcal{M}$ that describes the global properties of the training data. Note that this formulation of the statistical learning stage (also illustrated in Figure 1) can describe a discriminative function that requires features from the whole training set to compute separation boundaries between classes. In the case of generative learning, the output of the function $\mathcal{S}$ can be separated into Q independent models $\{\mathcal{M}_q\}$ containing parameters for each category, or into M independent models $\{\mathcal{M}_m\}$ corresponding to each training signal.

Once the training phase has been completed and a model $\mathcal{M}$ has been learned, the transform $\mathcal{T}$ is applied in the test phase to a new unlabeled recording s_{new}, leading to a sequence of features x_{new}. A function $\mathcal{G} : \mathcal{G}(x_{new}, \mathcal{M}) = c_{new}$ is then employed to classify the signal, returning a label in the set $\{\gamma_q\}_{q=1}^{Q}$.

Most of the algorithms mentioned in the section "Background: A History of ASC" follow the framework depicted in Figure 1 and only differ in their choice of the functions $\mathcal{T}, \mathcal{S}$, and $\mathcal{G}$. Some follow a seemingly different strategy but can still

be analyzed in light of this framework. For example, matrix factorizations algorithms like the one proposed by Benetos et al. [6] can be interpreted as combining features extraction and statistical modeling through the unsupervised learning of spectral templates and an activation matrix, as already discussed in the section "Features."

A special case of ASC framework is the so-called bag-of-frames approach [2], named in an analogy with the bag-of-words technique for text classification whereby documents are described by the distribution of their word occurrences. Bag-of-frames techniques follow the general structure shown in Figure 1 but ignore the ordering of the sequence of features when learning statistical models.

CHALLENGE OF DCASE

Despite a rich literature on systems for ASC, the research community has so far lacked a coordinated effort to evaluate and benchmark algorithms that tackle this problem. The challenge of DCASE has been organized in partnership with the IEEE Audio and Acoustic Signal Processing (AASP) Technical Committee to test and compare algorithms for ASC and for event detection and classification. This initiative is in line with a wider trend in the signal processing community aimed at promoting reproducible research [50]. Similar challenges have been organized in the areas of music information retrieval [36], speech recognition [5], and source separation [46].

THE DCASE DATA SET

Existing algorithms for ASC have been generally tested on data sets that are not publicly available [42], [20], making it difficult if not impossible to produce sustainable and reproducible experiments built on previous research. Creative Commons licensed sounds can be accessed for research purposes on http://freesound.org, a collaborative database that includes environmental sounds along with music, speech, and audio effects. However, the different recording conditions and varying quality of the data present in this repository would require a substantial curating effort to identify a set of signals suited for a rigorous and fair evaluation of ASC systems. On the other hand, the adoption of commercially available databases, such as the Series 6000 General Sound Effects Library [54], would constitute a barrier to research reproducibility due to their purchase cost.

The DCASE challenge data set [23] was specially created to provide researchers with a standardized set of recordings produced in ten different urban environments. The soundscapes were recorded in the London area and include: a bus, a busy street, an office, an open-air market, a park, a quiet street, a restaurant, a supermarket, the tube (underground railway), and a tube station. Two disjoint data sets were constructed from the same group of recordings, each containing ten 30-s long clips for each scene, totaling 100 recordings. Of these two data sets, one is publicly available and can be used by researchers to train and test their ASC algorithms; the other has been held back and has been used to evaluate the methods submitted for the challenge.

LIST OF SUBMISSIONS

A total of 11 algorithms were proposed for the DCASE challenge on ASC from research institutions worldwide. The respective authors submitted accompanying extended abstracts describing their techniques, which can be accessed from the DCASE Web site [55]. Table 2 lists the authors and titles of the contributions and defines the acronyms that are used throughout the article to refer to the algorithms.

In addition to the methods submitted for the challenge, we designed a benchmark baseline system that employs MFCCs, GMMs, and a maximum likelihood criterion. We have chosen to use these components because they represent standard practices in audio analysis, which are not specifically tailored to the ASC problem and, therefore, provide an interesting comparison with more sophisticated techniques. Table 1 summarizes the various approaches for ASC (see "MFCCs, GMMs, and a Maximum Likelihood Criterion").

EVALUATION OF ALGORITHMS FOR ASC

EXPERIMENTAL DESIGN

A system designed for ASC comprises training and test phases. The researchers who participated in the DCASE challenge were provided with a public data set that includes ground truth labels, indicating the environment in which the sounds were recorded. The training, test, and optimization of design parameters can be performed by partitioning this data set into training and test subsets, a standard practice in machine learning that is further discussed next. To obtain a fair evaluation reflecting the conditions of a real-world application, where sounds and labels are unknown to the

algorithms, the methods submitted to the DCASE challenge were tested on a private data set.

CROSS-VALIDATION

Recall from Figure 1 that statistical models are learned from the elements of the training data that belong to different classes and, therefore, depend on the particular signals available for training. This represents a general problem of statistical inference occurring every time models are learned using a limited set of data and is associated with a sampling error or bias. For example, to learn a statistical model of the sounds produced in an office environment, we would ideally need complete and continuous historical recordings from every office in the world. By only analyzing data recorded from one or several offices, we are bound to learn models that are biased toward the sounds present within the available signals. However, if the training data are rich enough to include sounds produced in most office environments, and if these sounds are effectively modeled, then the sampling bias can be bounded and models can statistically infer general properties of office environments from an incomplete set of measurements. Cross-validation is employed to minimize the sampling bias by optimizing the use of a set of available data. The collection of labeled recordings is partitioned into different subsets for training and testing so that all of the samples are used in the test phase. Different partition methods have been proposed in the literature for this purpose [7]. To evaluate the algorithms submitted to the DCASE challenge, we employed a so-called stratified fivefold cross-validation of the private data set. From 100 available recordings, five independent classifications are performed so that each run contains 80 training recordings and 20 test recordings. The
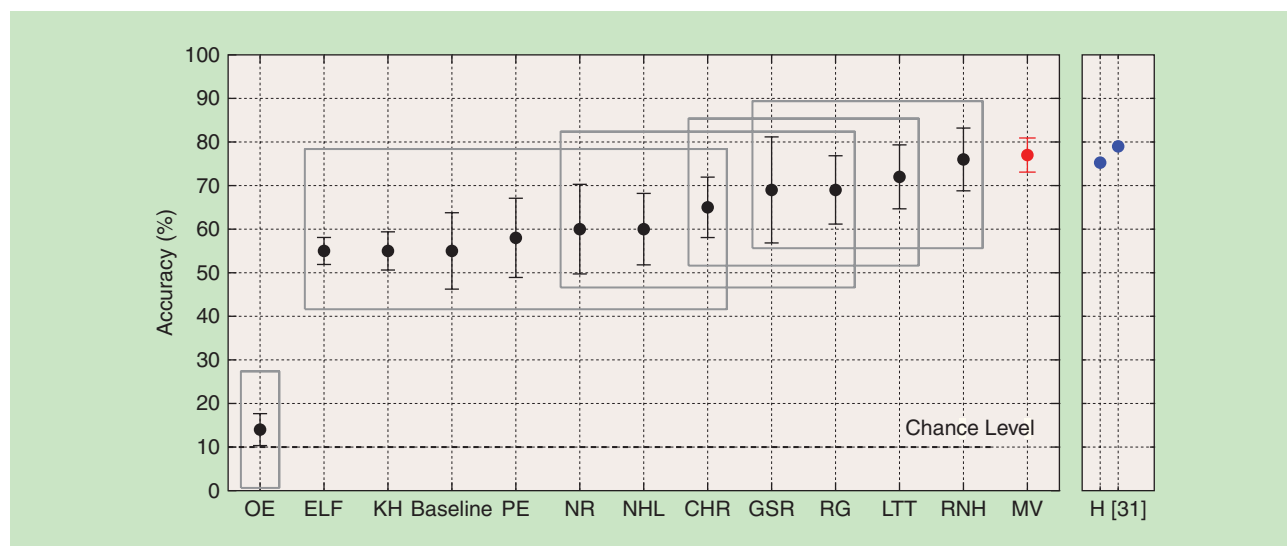
[FIG2] The mean values and confidence intervals of the accuracy of methods for ASC evaluated on the DCASE private data set using stratified fivefold cross-validation. The boxes enclose methods that cannot be judged to perform differently with a significance level of 95%. See Table 1 for the definition of the algorithms' acronyms. MV is a majority vote classifier that assigns to an audio recording the label that is most commonly returned by the other methods. H indicates the median human accuracy, as obtained through the test described in the section "Human Listening Test," while [31] refers to the human accuracy obtained by Krijnders and Holt. Note that algorithmic results are not directly comparable to the variations in human performance, and, hence, only the median human performance is depicted. See Figure 6 for more details on the distribution of human accuracies.

Estimate

Category	Bus	Busy Street	Office	Open-Air Market	Park	Quiet Street	Restaurant	Supermarket	Tube	Tube Station
Bus	100									
Busy Street		100								
Office			90			10				
Open-Air Market				80			10	10		
Park			10		80	10				
Quiet Street						50	50			
Restaurant				20			70	10		
Supermarket	10		10				10	60		10
Tube	10								80	10
Tube Station				10					30	60
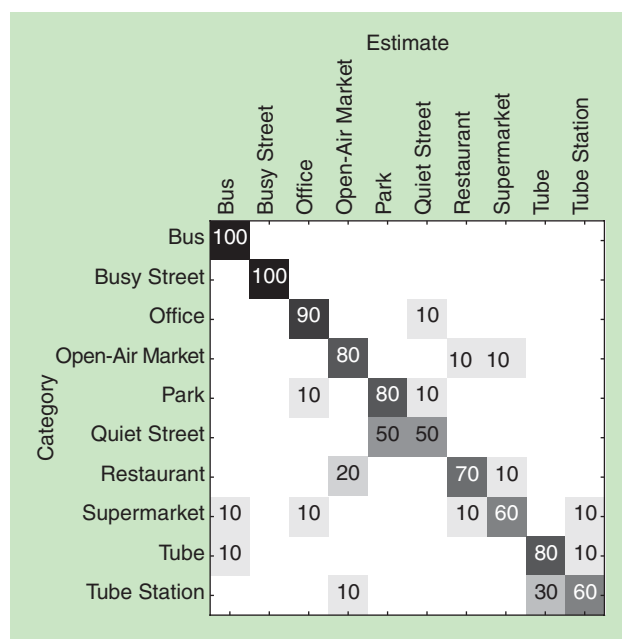
[FIG3] A confusion matrix of MV algorithmic classification results.

partitions are designed so that the five test subsets are disjoint, thus allowing for the classification of each of the 100 signals in the test phases. In addition, the proportion of signals belonging to different classes is kept constant in each training and test subset (eight signals per class in the former and two signals per class in the latter) to avoid class biases during the statistical learning.

PERFORMANCE METRICS

Performance metrics were calculated from each classification obtained using the training and test subsets, yielding five results for each algorithm. Let Γ be the set of correctly classified samples. The classification accuracy is defined as the proportion of correctly classified sounds relative to the total number of test samples. The confusion matrix is a $Q \times Q$ matrix whose (i, j)th element indicates the number of elements belonging to the ith class that have been classified as belonging to the jth class. In a problem with $Q = 10$ different classes, chance classification has an accuracy of 0.1 and a perfect classifier has an accuracy of 1. The confusion matrix of a perfect classifier is a diagonal matrix whose (i, i)th elements correspond to the number of samples belonging to the ith class.

RESULTS

Figure 2 depicts the results for the algorithms submitted to the DCASE challenge (see Table 1 for the acronyms of the methods). The central dots are the percentage accuracies of each technique calculated by averaging the results obtained from the five folds, and the bars are the relative confidence intervals. These intervals are defined by assuming that the accuracy value obtained from each fold is a realization of a Gaussian process whose expectation is the true value of the overall accuracy (i.e., the value that we would be able to measure if we evaluated

an infinite number of folds). The total length of each bar is the magnitude of a symmetric confidence interval computed as the product of the 95% quantile of a standard normal distribution $q_{N(0,1)}^{0.95} \approx 3.92$ and the standard error of the accuracy (that is, the ratio between the standard deviation of the accuracies of the folds and the square root of the number of folds $\sigma/\sqrt{5}$). Under the Gaussian assumption, confidence intervals are interpreted as covering with 95% probability the true value of the expectation of the accuracy.

From analyzing the plot, we can observe that the baseline algorithm achieves a mean accuracy of 55%, and a group of other methods obtain a similar result in the range between 55 and 65%. Four algorithms (GSR, RG, LTT, and RNH) approach or exceed a mean accuracy of 70%. OE performs relatively close to chance level and significantly worse than all of the other methods. The boxes displaying the results of the paired tests explained in the section "Ranking of Algorithms" indicate that a number of systems performed significantly better than the baseline.

Finally, the method MV indicated in red refers to a majority vote classifier whose output for each test file is the most common category assigned by all other methods. The mean accuracy obtained with this metaheuristic outperforms all of the other techniques, indicating a certain degree of independence between the classification errors committed by the algorithms. In other words, for almost 80% of soundscapes, some algorithms make a correct decision, and the algorithms that make an incorrect classification do not all agree on one particular incorrect label. This allows the decisions to be combined into a relatively robust metaclassifier. On the other hand, the performance obtained using MV is still far from perfect, suggesting that a number of acoustic scenes are misclassified by most algorithms. Indeed, this can be confirmed by analyzing the confusion matrix of the MV solution. As we can see in Figure 3, the class pairs (park, quiet street) and (tube, tube station) are commonly misclassified by the majority of the algorithms.

To investigate the poor performance of the method OE, we considered the results obtained on the public DCASE data set, which are not detailed here for the sake of conciseness. OE obtained the highest classification accuracy of all methods, suggesting that it overfitted the training data by learning models that could not generalize to the test signals.

RANKING OF ALGORITHMS

The ASC performance has been evaluated by computing the statistics among different cross-validation folds. However, all of the submitted methods have been tested on every file of the same held-back data set, and this allows us to compare their accuracy on a file-by-file basis. Recall that s_p indicates a signal in the test set. A binary variable X_p can be assigned to each signal and defined so that it takes the value 1 if the file has been correctly classified and 0 if it has been misclassified. Each X_p can thus be interpreted as a realization of a Bernoulli random process whose average is the mean accuracy of the classifier.

Given two classifiers C_1, C_2 and the corresponding variables $X_{C1,p}, X_{C2,p}$, a third random variable $Y_p = X_{C1,p} - X_{C2,p}$ assumes
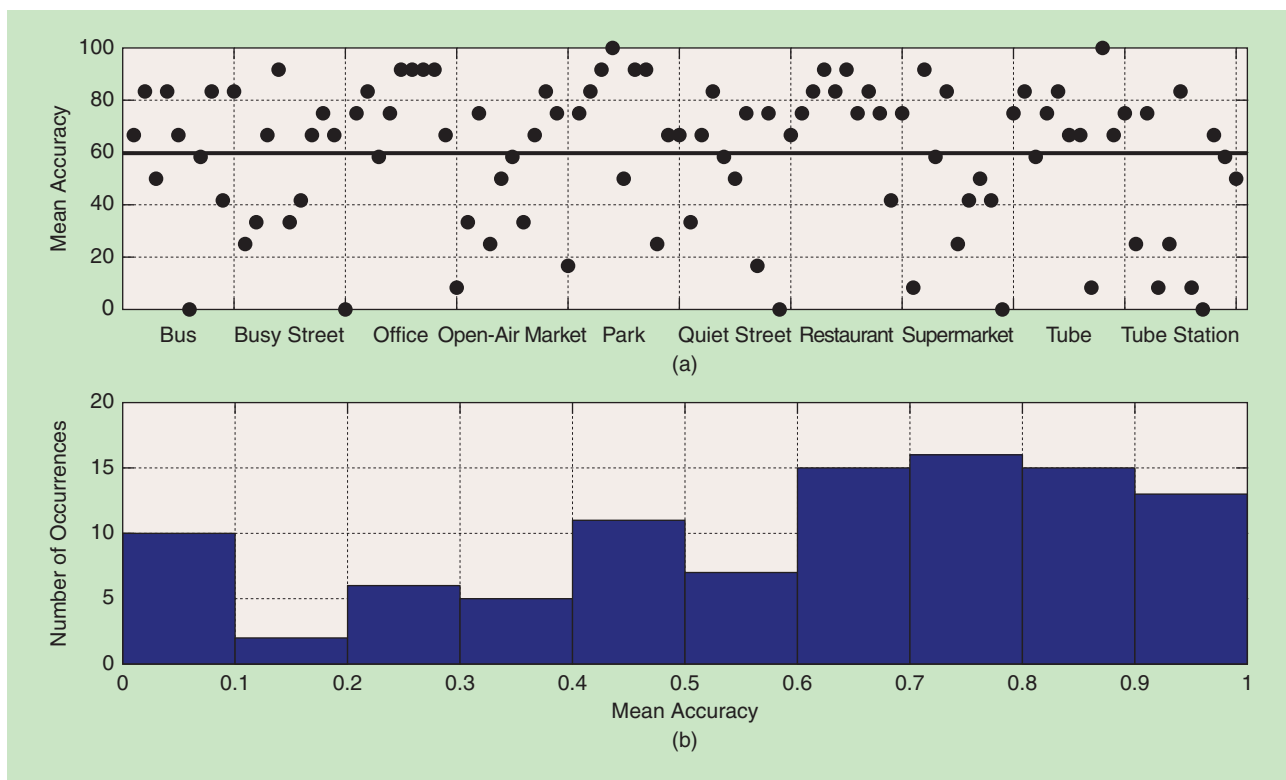
[FIG4] The distribution of algorithmic soundscapes classification accuracies. The solid line in (a) represents the average accuracy calculated from all of the acoustic scenes. (b) The histogram of mean accuracies resulting from the classification of all 100 soundscapes, highlighting that ten soundscapes are correctly classified by at most only 10% of the algorithms.

values in the set $\{-1, 0, +1\}$ and indicates the difference in the correct or incorrect classification of s_p by the two classifiers (that is, $Y = -1$ implies that C_1 has misclassified s and C_2 has correctly classified it; $Y = 0$ means that the two methods return equivalently correct or incorrect decisions, and $Y = 1$ implies that C_1 has correctly classified s and C_2 has misclassified it). A sign test [24] can be performed to test the hypothesis that the expected value of Y is equal to zero. This is equivalent to performing a paired test evaluating the hypothesis that the performance of the two classifiers C_1 and C_2 is the same. Hence, being able to reject this hypothesis at a fixed probability level provides a method to rank the algorithms.

The gray boxes in Figure 2 represent groups of methods whose accuracy is not significantly different when tested on the DCASE data set, according to the sign tests ranking criterion evaluated between pairs of different methods. Methods enclosed in the same box cannot be judged to perform better or worse according to the chosen significance level. Starting with the least accurate algorithms, we can observe that the performance of OE is significantly different compared with all the other techniques. Then, a cluster of methods ranging from ELF to CHR do not perform significantly differently from the baseline. GSR and RG can be said to have significantly higher accuracy if compared to the baseline method, but not if compared to NR, NHL, or CHR. Finally, RNH is not significantly more accurate than GSR, RG, and LTT, but it outperforms all of the remaining methods. Note

that we do not include the results of the majority vote metaheuristic in the ranking, as a paired sign test assumes the variables $X_{C1,p}$, $X_{C2,p}$ to be statistically independent, and this assumption is violated in the case of MV.
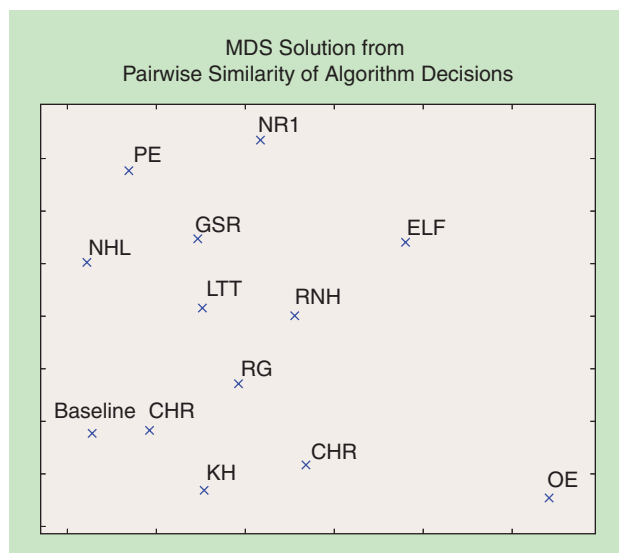
[FIG5] A multidimensional scaling solution (2-D) derived from the pairwise similarities between algorithm labeling decisions. Algorithms that make similar (mis)classifications tend to appear close to one another.
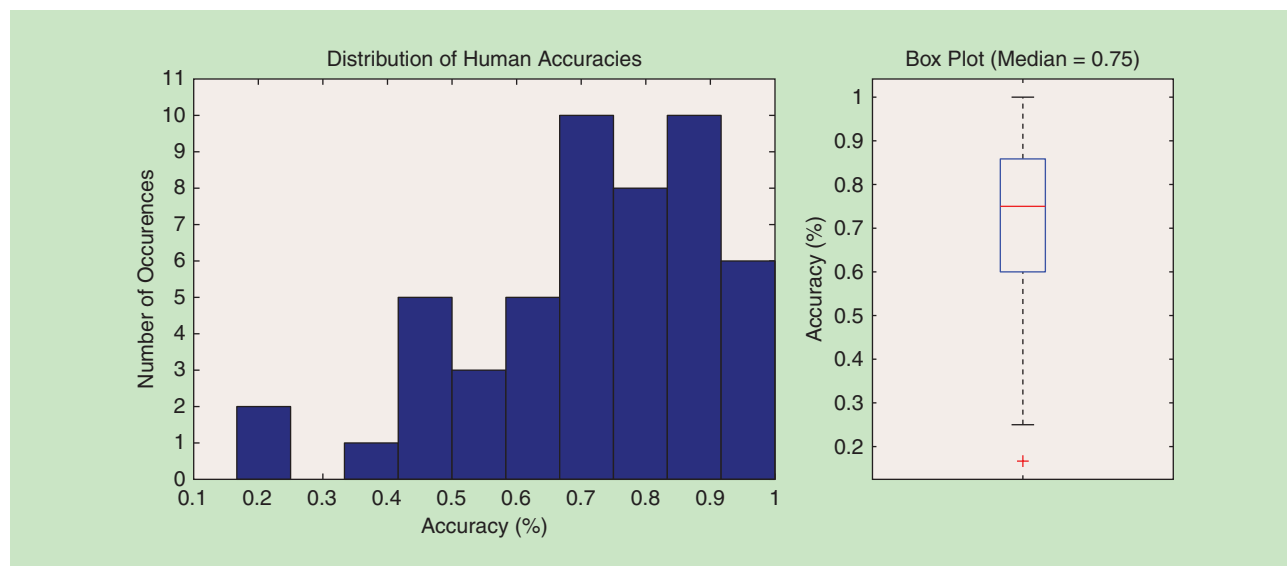
[FIG6] The distribution of human soundscape classification accuracies.

DISTRIBUTION OF ALGORITHMIC SOUNDSCAPES CLASSIFICATION ACCURACIES

Further analysis of the classification results can be carried out to understand whether there are individual soundscape recordings in the DCASE data set that are classified more accurately than others. After evaluating each method with a fivefold cross-validation, every signal s_p is classified by all of the algorithms. Figure 4 shows a scatter plot of the mean classification accuracy obtained for each file and a histogram of the relative distribution. We can observe that some acoustic scenes belonging to the categories "bus," "busy street," "quiet street," and "tube station" are never correctly classified (those at 0%). In general, the classification accuracy among soundscapes belonging to the same category greatly varies, with the exception of the classes "office" and "restaurant" that might contain distinctive events or sound characteristics resulting in more consistent classification accuracies.

PAIRWISE SIMILARITY OF ALGORITHMS' DECISIONS

While the results in Figure 2 demonstrate the overall accuracy achieved by algorithms, they do not show which algorithms tend to make the same decisions as others. For example, if two algorithms use a very similar method, we would expect them to make a similar pattern of mistakes. We can explore this aspect of the algorithms by comparing their decisions pairwise against one another and using the number of disagreements as a distance measure. We can then visualize this using multidimensional scaling (MDS) to project the points into a low-dimensional space, which approximately honors the distance values [16, Ch. 10].

The results of the MDS are shown in Figure 5. We tested multiple dimensionalities and found that 2-D (as shown) yielded a sufficiently low stress to be suitably representative. The OE submission is placed in a corner of the plot at some distance from the other algorithms; that submission achieved low scores on the private testing data. As a whole, the plot does not appear to cluster together methods by feature type, as MFCC and non-MFCC approaches as well as SVM and non-SVM approaches are interspersed.

HUMAN LISTENING TEST

To determine a human benchmark for the algorithmic results on ASC, we designed a crowdsourced online listening test in which participants were asked to classify the public DCASE data set by listening to the audio signals and choosing the environment in which each signal was recorded from the ten categories: "bus," "busy street," "office," "open-air market," "park," "quiet street," "restaurant," "supermarket," "tube," and "tube station."

In designing the listening experiment, we chose not to divide the classification into training and test phases because we were interested in evaluating how well humans can recognize the acoustic environments basing their judgment on nothing other than their own personal experience. The participants were not presented with labeled training sounds before the test, and they were not told their performance during the test.

To maximize the number of people taking the test, we allowed each participant to classify as many acoustic scenes as he or she wanted while randomizing the order in which the audio samples appeared in the test to ensure that each file had the same probability of being classified. To avoid potential biases, people who were likely to have worked with the data and, thus, were likely to know the class labels in advance, did not take the test.

HUMAN ACCURACY

Fifty participants took part in the test. Their most common age was between 25 and 34 years old, and the most common listening device employed during the test was high-quality headphones. Special care was taken to remove test cases or invalid attempts from the sample. This included participants clearly labeled as "test" in the metadata, participants who only attempted to label only one or

two soundscapes, and most of those who achieved scores as low as 0%, which points to outliers with a clear lack of motivation. Figure 6 shows that the mean accuracy among all participants was 72%, and the distribution of accuracies reveals that most people scored between 60 and 100%, with two outliers whose accuracy was as low as 20%. Since the distribution of accuracies is not symmetric, we show a box plot summarizing its statistics instead of reporting confidence intervals for the mean accuracy. The median value of the participants' accuracy was 75%, the first and third quartiles are located at around 60 and 85%, and 95% of values lie between around 45 and 100%. Note that, although we decided to include the results from all of the participants in the study who classified at least a few soundscapes, the most extreme points (corresponding to individuals who obtained accuracies of about 25 and 100%, respectively) only include classifications performed on fewer than ten acoustic scenes. Removing from the results participants who achieved about 25% accuracy would result in a mean of 74%, which is a lot closer to the median value. In a more controlled listening test, Krijnders and Holt [31] engaged 37 participants, with each participant asked to listen to 50 public DCASE soundscapes and select one of the ten categories. The participants were required to listen for the entire duration of the recordings and use the same listening device. They obtained a mean accuracy of 79%, which is in the same area as the results of our crowdsourced study (75%).

CUMULATIVE ACCURACY
During the test, we asked the participants to indicate their age and the device they used to listen to the audio signals, but we did not observe a correlation between these variables and the classification accuracy. We did observe a correlation between the number of classified samples and the overall classification accuracy. People who listened to and categorized most or all of the 100 total samples tended to score better than individuals who only classified a few sounds. To assess whether this occurred because participants learned how to better classify the sounds as they progressed in the test, we computed for each individual the cumulative accuracy $\rho(t)$, which is defined as the ratio between the number of correctly classified samples and the total number of classified samples at times $t = 1, \ldots, P$

$$\rho(t) = \frac{|\Gamma(t)|}{t}. \tag{3}$$

A positive value of the discrete first-time derivative of this function $\rho'(t) = \rho(t) - \rho(t-1)$ would indicate that there is an improvement in the cumulative classification accuracy as time progresses. Therefore, we can study the distribution of $\rho'(t)$ to assess the hypothesis that participants were implicitly training an internal model of the classes as they performed the test. The average of the function $\rho'(t)$ calculated for all of the participants was -0.0028. A right-tailed t-test rejected with 95% probability that the expectation of $\rho'(t)$ is greater than zero, and a left-tailed t-test failed to reject with the same probability that the expectation is less than zero, indicating that participants did not improve their accuracy as they progressed through the test. This is a positive finding as the listening test was designed to avoid training from

[FIG7] A confusion matrix of human classification results. Note that the rows of the confusion matrix might not add up to 100% due to the rounding of percentages.

the exposure to the soundscapes. Having rejected the learning hypothesis, we are left with a selection bias explanation: we believe that people who classified more sounds were simply better able or more motivated to do the test than individuals who found the questions difficult or tedious and did not perform as well.

SCENES CLASS CONFUSION MATRIX
Further insight about the human classification results can be obtained by analyzing the overall confusion matrix of the listening test. Figure 7 shows that "supermarket" and "open-air market" are the most commonly misclassified categories whose samples have been estimated as belonging to various other classes. In addition, there are some common misclassifications between the classes "park" and "quiet street" and, to a minor extent, between the classes "tube" and "tube station."

DISTRIBUTION OF HUMAN SOUNDSCAPES CLASSIFICATION ACCURACIES
To assess if some soundscapes were classified more accurately than others, we conducted a similar analysis for the human performance benchmark to the one described in the section "Distribution of Algorithmic Soundscapes Classification Accuracies." Figure 8 depicts the mean accuracy of the classification of 100 soundscapes in the public DCASE data set and a histogram of the relative distribution. The public and private portions of the DCASE data set are disjoint subsets of the group of recordings produced for the challenge; therefore, a paired comparison of the accuracies in Figures 4 and 8 cannot be carried out. Nonetheless, it is informative to compare the trends between the two analyses: it appears that the mean performance for the human classification approaches 80% as opposed to a value of around 55%
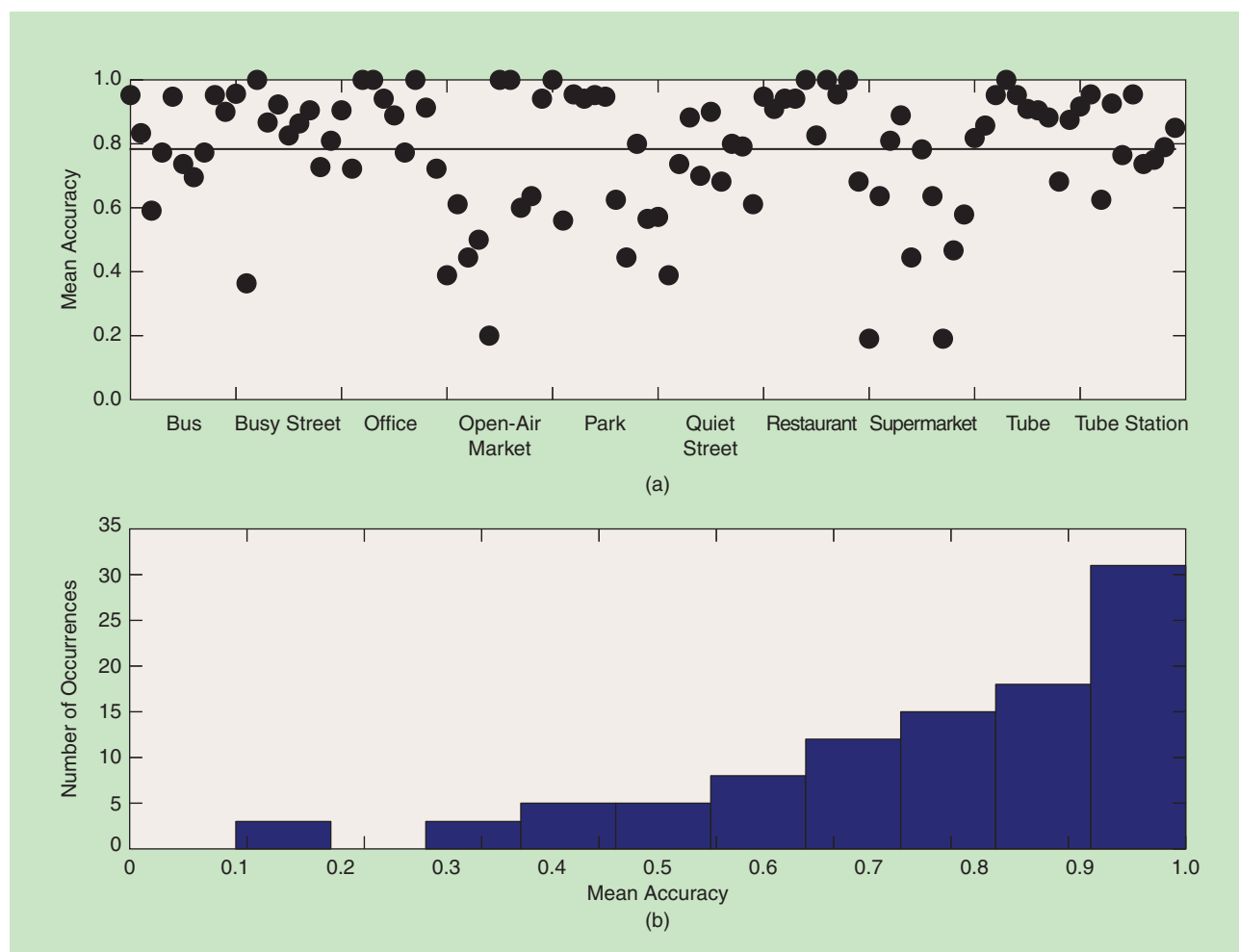
[FIG8] The distribution of human soundscapes classification accuracies. (a) Each point represents the mean classification accuracy for a given acoustic scene, and the solid line represents the mean accuracy across all scenes. (b) The histogram depicts the distribution of classification accuracies across acoustic scenes.

achieved on average by the algorithms. In addition, the distribution of the mean accuracy in the case of human classification appears more regular, with most soundscapes being correctly classified most of the time, and with only a few outlier scenes whose classification accuracy is below 30%.

DISCUSSION

By interpreting sophisticated algorithms in terms of a general framework, we have offered a tutorial that uncovers the most important factors to take into account when tackling a difficult machine-learning task such as the classification of soundscapes. Inevitably, every abstraction or generalization is carried out at the expense of omissions in the description of the implementation details of each method. Nonetheless, we think that valuable insights can be gained by analyzing the classification results in light of the framework proposed in the section "A General Framework for ASC."

ALGORITHMS FROM THE DCASE CHALLENGE

A first trend regarding the choice of statistical learning function S can be inferred by analyzing the algorithms submitted for the

DCASE challenge summarized in Table 2. All but one method (ELF) use discriminative learning to map features extracted from the audio signals s_m to class labels c_m. Moreover, most of the algorithms whose mean accuracy is greater than or equal to that achieved by the baseline method employ SVM. All techniques that perform significantly better than the baseline, except for LTT, employ a combination of generative and discriminative learning by training an SVM classifier using parameters of models $\mathcal{M}_m$ learned from individual audio scenes. This suggests that models learned from single audio scenes offer an appropriate tradeoff between discrimination and generalization. On the one hand, audio signals recorded in the same environment are analyzed by learning different statistical models that account for variations between one recording and the next. On the other hand, the parameters of these models occupy localized regions in a parameter's space so that classification boundaries can be learned to discriminate between signals recorded in different environments.

A closer analysis of some of the better-scoring algorithms (GSR, RG, and RNH) reveals a further common design motivation. In different ways, all three methods attempt to model

[TABLE 2] A SUMMARY AND CATEGORIZATION OF COMPUTATIONAL METHODS FOR ASC. THE ACRONYMS AFTER THE AUTHOR(S) NAME(S) IN THE METHOD COLUMN ARE DEFINED IN TABLE 1. THE ARROWS INDICATE SEQUENTIAL PROCESSING, E.G., WHEN STATISTICAL PARAMETERS LEARNED FROM FEATURES ARE FED TO AN SVM TO OBTAIN SEPARATING HYPERPLANES. IN SOME CASES, THE DECISION CRITERION OF SVMs (ONE VERSUS ALL, ONE VERSUS ONE, OR ALTERNATIVE) IS NOT SPECIFIED IN THE REFERENCE. HOWEVER, IT IS ALWAYS SPECIFIED WHEN THE DISCRIMINATIVE LEARNING IS PERFORMED ON FRAMES, AND AN OVERALL CLASSIFICATION IS DETERMINED BY A MAJORITY VOTE OR A WEIGHTED MAJORITY VOTE. NOTE THAT, FOR EACH WORK CITED, ONLY THE METHOD LEADING TO THE BEST CLASSIFICATION RESULTS WAS CONSIDERED.

METHOD	FEATURES	STATISTICAL MODEL	DECISION CRITERION
SAWHNEY AND MAES [42]	FILTER BANK	NONE	NEAREST NEIGHBOR → MAJORITY VOTE
CLARKSON ET AL. [12]	MFCs	HMM	MAXIMUM LIKELIHOOD
ERONEN ET AL. [20]	MFCCs, LOW-LEVEL DESCRIPTORS, ENERGY/FREQUENCY, LPCS → ICA, PCA	DISCRIMINATIVE HMM	MAXIMUM LIKELIHOOD
AUCOUTURIER [2]	MFCCs	GMMs	NEAREST NEIGHBOR
CHU ET AL. [10]	MFCCs, PARAMETRIC (GABOR)	GMMs	MAXIMUM LIKELIHOOD
MALKIN AND WAIBEL [34]	MFCCs, LOW-LEVEL DESCRIPTORS → PCA	LINEAR AUTOENCODER NETWORKS	MAXIMUM LIKELIHOOD
CAUCHI [8]	NMF		MAXIMUM LIKELIHOOD
BENETOS [6]	PLCA		MAXIMUM LIKELIHOOD
HEITTOLA ET AL. [27]	ACOUSTIC EVENTS	HISTOGRAM	MAXIMUM LIKELIHOOD
CHAUDHURI ET AL. [9]	AUDs	N-GRAM LANGUAGE MODELS	MAXIMUM LIKELIHOOD
DCASE SUBMISSIONS			
BASELINE	MFCCs	GMMs	MAXIMUM LIKELIHOOD
RNH	MFCCs	RQA, MOMENTS → SVM	—
RG	LOCAL GRADIENT HISTOGRAMS (LEARNED ON TIME–FREQUENCY PATCHES)	AGGREGATION → SVM	ONE VERSUS ONE
GSR	MFCCs, ENERGY/FREQUENCY, VOICING	MOMENTS, PERCENTILES, LINEAR REGRESSION COEFFICIENT → SVM	MAJORITY VOTE
CHR	ENERGY/FREQUENCY	SVM	ONE VERSUS ALL, MAJORITY VOTE
NHL	LEARNED (MFCCs → SRBM)	SELECTIVE MAX POOLING → SVM	ONE VERSUS ALL
NR	MFCCs, ENERGY/FREQUENCY, SPATIAL → FISHER FEATURE SELECTION	SVM	MAJORITY VOTE
PE	FILTER BANK → PARAMETRIC (GABOR) → PCA	SVM	ONE VERSUS ONE, WEIGHTED MAJORITY VOTE
KH	→ VOICING	MOMENTS, PERCENTILES → SVM	—
ELF	MFCCs	I-VECTOR → $_P$LDA	MAXIMUM LIKELIHOOD
LTT	MFCCs	ENSEMBLE OF CLASSIFICATION TREES	MAJORITY VOTE → TREE BAGGER
OE	SIZE OF COMPRESSED AUDIO	COMPRESSION DISTANCE → ENSEMBLE OF CLASSIFICATION TREES	→ RANDOM FOREST

temporal relationships between features extracted from different portions of the signals. RNH employs RQA parameters to encode periodicities (or stationarity) of the MFCC coefficients, RG accounts for time–frequency structures in the audio signals by learning gradient histograms of images derived from their spectrograms, and, finally, GSR computes linear regression coefficients of local features that encode general trends across a whole scene. This supports the intuitive observation that an ASC method should take into consideration the time evolution of different acoustic events to model complex acoustic scenes.

A further observation derived from analyzing Table 2 is that, among the methods that used classification trees in combination with a tree bagger or a random forest algorithm, OE achieved a poor classification performance, while LTT reached the second-best mean accuracy. This might suggest that meta-algorithms can be a valuable strategy but may also be prone to overfitting.

Finally, a more exploratory remark regards the general use of the framework described in the section "A General Framework for ASC." Aucoutourier [3] studied the performance of a class of algorithms for audio timbre similarity, which followed a method similar to the ASC baseline. He reported the existence of a glass ceiling as more and more sophisticated algorithms failed to improve the performance obtained using a simple combination of MFCCs and GMMs. To a certain extent, the fact that seven out of 11 ASC methods did not significantly outperform our baseline might suggest a similar effect and

urge researchers to pursue alternative paradigms. Modeling temporal relationships as described before is one first step in this direction, and perhaps algorithms whose design motivations depart from those driving the development of the baseline, such as the normalized compression dissimilarity (OE), might be worth additional investigation.

COMPARISON OF HUMAN AND ALGORITHMIC RESULTS

When designing the human listening test, we chose to present individuals with samples from the public DCASE data set to avoid distributing the held-back data set that was produced to test the algorithms. In addition, we chose not to divide the human task into training and testing phases because we were interested in evaluating how people performed by only drawing from previous experience and not from prior knowledge about the test set. The different experimental design choices between human and algorithmic experiments do not allow us to perform a statistically rigorous comparison of the classification performances. However, since the public and private DCASE data sets are two parts of a unique session of recordings realized with the same equipment and in the same conditions, we still believe that qualitative comparisons are likely to reflect what the results would have been had we employed a different design strategy that allowed for a direct comparison. More importantly, we believe that qualitative conclusions about how well algorithms can approach human capabilities are more interesting than rigorous significance tests on how humans can perform according to protocols (such as the fivefold stratified cross-validation) that are clearly unnatural tasks.

Having specified the above disclaimer, several observations can be derived from comparing algorithmic and human classification results. First, Figures 2 and 6 show that RNH achieves a mean accuracy in the classification of soundscapes of the private DCASE data set that is similar to the median accuracy obtained by humans on the public DCASE data set. This strongly suggests that the best-performing algorithm achieves similar accuracy compared to a median human benchmark.

Second, the analysis of the misclassified acoustic scenes summarized in Figures 4 and 8 suggests that, by aggregating the results from all of the individuals who took part in the listening test, all of the acoustic scenes are correctly classified by at least some individuals, while there are scenes that are misclassified by all algorithms. This observation echoes the problem of hubs encountered in music information retrieval, whereby certain songs are always misclassified by algorithms [41]. Moreover, unlike for the algorithmic results, the distribution of human errors shows a gradual decrease in accuracy from the easiest to the most challenging soundscapes. This observation indicates that, in the aggregate, the knowledge acquired by humans through experience still results in a better classification of soundscapes that might be considered ambiguous or lacking in highly distinctive elements.

Finally, the comparison of the confusion matrices presented in Figures 3 and 7 reveals that similar pairs of classes (such as "park" and "quiet street" or "tube" and "tube station") are commonly misclassified by both humans and algorithms. Given what we found about the misclassification of a single acoustic scene, we do not infer from this observation that the algorithms are using techniques that emulate human audition. An alternative interpretation is rather that some groups of classes are inherently more ambiguous than others because they contain similar sound events. Even if both physical and semantic boundaries between environments can be inherently ambiguous, for the purpose of training a classifier, the universe of soundscapes classes should be defined as mutually exclusive and collectively exhaustive. In other words, it should include all of the possible categories relevant to an ASC application while ensuring that every category is as distinct as possible from all of the others.

FURTHER RESEARCH

Some themes that have not been considered in this article may be important depending on particular ASC applications and are suggested here for further research.

1) *Algorithm complexity*: A first issue to be considered is the complexity of algorithms designed to learn and classify acoustic scenes. Given that mobile context-aware services are among the most relevant applications of ASC, particular emphasis should be placed on designing methods that can be $\mathcal{G}$ run with the limited processing power available to smartphones and tablets. The resources-intensive processing of training signals to learn statistical models for classification can be carried out off-line, but the operators $\mathcal{T}$ and $\mathcal{G}$ still need to be applied to unlabeled signals and, depending on the application, might need to be simple enough to allow real-time classification results.

2) *Continuous and user-assisted learning*: Instead of assuming a fixed set of categories, as done in most publications on ASC, a system might be designed to be progressively trained to recognize different environments. In this case, a user should record soundscape examples that are used to train classification models (either online or off-line, using the recording device's own computational resources or uploading and processing the signals with remote cloud resources) and progressively add new categories to the system's memory of soundscapes. Users could also assist the training by confirming or rejecting the category returned from querying each unlabeled signal and, thus, refine the statistical models every time a new classification is performed. Such systems would inevitably require more intervention by the user but would likely be more precise and relevant than totally automated systems.

3) *Hierarchical classification*: In this article, we have considered a set of categories whose elements are assumed to be mutually exclusive (that is, a soundscape can be classified as bus or park but not both). Alternatively, a hierarchical classification could be considered where certain categories are subsets or supersets of others. For example, a system might be designed to classify between outdoor and indoor environments and then distinguish between different subsets of the two general classes. In this context, different costs could be associated with different types of misclassification errors: for example, algorithms could be trained to be very accurate in discriminating between outdoor and indoor and less precise

in distinguishing between an outdoor park and an outdoor busy street.

4) *Acoustic scene detection*: As a limit case of systems that employs nonuniform misclassification costs, algorithms might be designed to detect a particular environment and group all of the other irrelevant categories into an others class. In this case, the system would essentially perform acoustic scene detection rather than classification.

5) *Multimodal learning*: Another avenue of future research consists in fusing multimodal information to improve the classification accuracy of ASC systems. Video recordings, geo-location information, and temperature and humidity sensors are all examples of data that can be used in conjunction with audio signals to provide machines with context awareness.

6) *Event detection and scene classification*: The combination of event detection algorithms and ASC, which has already been the object of research endeavors [27], [9], is likely to benefit from advances in both areas. Information regarding the events occurring in an acoustic scene could be combined with more traditional frame-based approaches to update the probability of categories as different events are detected. For example, while general spectral properties of a soundscape could be used to infer that a signal was likely to have been recorded in either a park or on a quiet street, detecting the event car horn would help disambiguate between the two. Furthermore, this Bayesian strategy employed to update the posterior probability of different classes could be used to handle transitions between different environments.

7) *Testing on different data sets*: Finally, data sets that contain sounds from different acoustic environments have been recently released. They include the diverse environments multichannel acoustic noise database [48] and the database of annotated real environmental sounds [49].

CONCLUSIONS

In this article, we provided a tutorial on ASC with a particular emphasis on computational algorithms designed to perform this task automatically. By introducing a framework for ASC, we have analyzed and compared methods proposed in the literature in terms of their modular components. We then presented the results of the DCASE challenge, which set the state of the art in computational ASC, and compared the results obtained by algorithms with a baseline method and a human benchmark. On the one hand, many of the submitted techniques failed to significantly outperform the baseline system, which was designed to not be optimized for this particular task. However, some methods significantly outperformed the baseline and approached an accuracy comparable to the human benchmark. Nonetheless, a more careful analysis of the human and algorithmic results highlighted that some acoustic scenes were misclassified by all algorithms while all soundscapes were correctly classified by at least some individuals. This suggests that there is still room for improvement before algorithms reach and surpass the ability of humans to make sense of their environment based on the sounds it produces.

AUTHORS

Daniele Barchiesi (d.barchiesi@ucl.ac.uk) received the M.Sc. degree in 2009 and the Ph.D. degree in 2013 both from the Centre for Digital Music at Queen Mary University of London. He has worked on research topics in the fields of audio engineering, machine learning, sparse approximation, and reproducible research. Since January 2014, he has been a research associate in data science at University College London, where he is investigating big data in the context of computational social science.

Dimitrios Giannoulis (Dimitrios.Giannoulis@eecs.qmul.ac.uk) received the B.Sc. degree in physics from the National University of Athens, Greece, in 2009. He received the M.Sc. and Ph.D. degrees in electronic engineering from Queen Mary University of London in 2010 and 2014, respectively. He is currently an assistant vice president risk modeler in counterparty credit risk at Credit Suisse, London. His research interests include signal processing and machine learning. He is a Senior Member of the IEEE.

Dan Stowell (dan.stowell@qmul.ac.uk) is a researcher applying machine learning to sound. He has worked on voice, music, birdsong, and environmental soundscapes. He was recently awarded an Engineering and Physical Sciences Research Council Early Career Research Fellowship based at Queen Mary University of London, where he is developing new techniques in structured machine listening for soundscapes with multiple birds.

Mark D. Plumbley (mark.plumbley@eecs.qmul.ac.uk) received the B.A. (honors) degree in electrical sciences and the Ph.D. degree in neural networks from the University of Cambridge, United Kingdom, in 1984 and 1991, respectively. From 1991 to 2001, he was a lecturer at King's College London. He moved to Queen Mary University of London in 2002, where he is now an Engineering and Physical Sciences Research Council leadership fellow and the director of the Centre for Digital Music. His research focuses on the automatic analysis of music and other sounds, including automatic music transcription, beat tracking, and acoustic scene analysis, using methods such as source separation and sparse representations. He is a past chair of the Independent Component Analysis Steering Committee and is a member of the IEEE Signal Processing Society Technical Committee on Audio and Acoustic Signal Processing. He is a Senior Member of the IEEE.

REFERENCES

[1] (2013). Extended abstracts, IEEE AASP challenge on detection and classification of acoustic scenes and events. [Online]. Available: http://c4dm.eecs.qmul.ac.uk/sceneseventschallenge/

[2] J.-J. Aucouturier, B. Defreville, and F. Pachet, "The bag-of-frames approach to audio pattern recognition: A sufficient model for urban soundscapes but not for polyphonic music," *J. Acoust. Soc. Am.*, vol. 112, no. 2, pp. 881–891, 2007.

[3] J.-J. Aucouturier and F. Pachet, "Improving timbre similarity: How high is the sky?," *J. Negative Results Speech Audio Sci.*, vol. 1, no. 1, 2004.

[4] J. Ballas, "Common factors in the identification of an assortment of brief everyday sounds," *J. Exp. Psychol.: Hum. Percept. Perform.*, vol. 19, no. 2, pp. 250–267, 1993.

[5] J. Barker, E. Vincent, N. Ma, H. Christensen, and P. Green, "The PASCAL CHiME speech separation and recognition challenge," *Comput. Speech Lang.*, vol. 27, no. 3, pp. 621–633, 2012.

[6] E. Benetos, M. Lagrange, and S. Dixon, "Characterisation of acoustic scenes using a temporally-constrained shift-invariant model," in *Proc. 15th Int. Conf. Digital Audio Effects (DAFx-12)*, 2012.

[7] C. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2007.

[8] B. Cauchi, "Non-negative matrix factorization applied to auditory scene classification," Master's thesis, ATIAM (UPMC/IRCAM/TELECOM ParisTech), 2011.

[9] S. Chaudhuri, M. Harvilla, and B. Raj, "Unsupervised learning of acoustic unit descriptors for audio content representation and classification," in *Proc. INTERSPEECH*, 2011, pp. 2265–2268.

[10] S. Chu, S. Narayanan, and C.-C. Jay Kuo, "Environmental sound recognition with time-frequency audio features," *IEEE Trans. Audio, Speech Lang. Processing*, vol. 17, no. 6, pp. 1142–1158, Aug. 2009.

[11] S. Chu, S. Narayanan, C.-C. Jay Kuo, and M. J. Matari, "Where am I? Scene recognition for mobile robots using audio features," in *Proc. IEEE Int. Conf. Multimedia and Expo.*, 2006, pp. 885–888.

[12] B. Clarkson, N. Sawhney, and A. Pentland, "Auditory context awareness via wearable computing," in *Proc. 1998 Workshop Perceptual User Interfaces (PUI'98)*, 1998.

[13] B. Defréville, P. Roy, C. Rosin, and F. Pachet, "Automatic recognition of urban sound sources," in *Proc. 120th Audio Engineering Society Convention*, 2006, number 6827.

[14] N. Dehak, R. Dehak, P. Kenny, N. Brummer, P. Ouellet, and P. Dumouchel, "Support vector machines versus fast scoring in the low-dimensional total variability space for speaker verification," in *Proc. INTERSPEECH*, 2009.

[15] D. Dubois, C. Guastavino, and M. Raimbault, "A cognitive approach to urban soundscapes: Using verbal data to access everyday life auditory categories," *Acta Acustica*, vol. 92, pp. 865–874, 2006.

[16] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley-Interscience, 2000.

[17] K. El-Maleh, A. Samouelian, and P. Kabal, "Frame level noise classification in mobile environments," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 1999, vol. 1, pp. 237–240.

[18] D. P. W. Ellis, "Prediction-driven computational auditory scene analysis," Ph.D. dissertation, Dept. Electr. Eng. Comput. Sci., Massachusetts Institute of Technology, 1996.

[19] D. P. W. Ellis and K. Lee, "Minimal-impact audio-based personal archives," in *Proc. Workshop Continuous Archiving and Recording of Personal Experiences*, 2004, pp. 39–47.

[20] A. J. Eronen, V. T. Peltonen, J. T. Tuomi, A. P. Klapuri, S. Fagerlund, T. Sorsa, G. Lorho, and J. Huopaniemi, "Audio-based context recognition," *IEEE Trans. Audio, Speech Lang. Processing*, vol. 14, no. 1, pp. 321–329, Jan. 2006.

[21] A. J. Eronen, J. T. Tuomi, A. Klapuri, and S. Fagerlund, "Audio-based context awareness—Acoustic modeling and perceptual evaluation," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2003, vol. 5, pp. 529–532.

[22] P. Gaunard, C. G. Mubikangiey, C. Couvreur, and V. Fontaine, "Automatic classification of environmental noise events by hidden Markov models," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP '98)*, 1998, vol. 6, pp. 3609–3612.

[23] D. Giannoulis, D. Stowell, E. Benetos, M. Rossignol, M. Lagrange, and M. D. Plumbley, "A database and challenge for acoustic scene classification and event detection," in *Proc. European Signal Processing Conf.* (*EUSIPCO*), 2013.

[24] J. D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference,* 5th ed. London, U.K.: Chapman & Hall, 2010, Number 978-1420077612.

[25] J. M. Grey and J. W. Gordon, "Perceptual effects of spectral modifications on musical timbres," *J. Acoust. Soc. Am.*, vol. 63, no. 5, pp. 1493–1500, 1978.

[26] P. Guyot, J. Pinquier, and R. André-Obrecht, "Water sounds recognition based on physical models," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 793–797.

[27] T. Heittola, A. Mesaros, A. J. Eronen, and T. Virtanen, "Audio context recognition using audio event histogram," in *Proc. European Signal Processing Conf.* (*EUSIPCO*), 2010.

[28] T. Heittola, A. Mesaros, A. J. Eronen, and T. Virtanen, "Context-dependent sound event detection," *EURASIP J. Audio, Speech, Music Process.*, vol. 1, no. 1, pp. 1–13, Jan. 2013.

[29] G. Hu and D. Wang, "Auditory segmentation based on onset and offset analysis," *IEEE Trans. Audio, Speech, Lang. Processing,* vol. 15, no. 2, pp. 396–405, 2007.

[30] S. Ioffe, "Probabilistic linear discriminant analysis," in *Proc. 9th European Conf. Computer Vision* (*ECCV*), 2006, pp. 531–542.

[31] J. Krijnders and G. A. t. Holt, "Tone-fit and MFCC scene classification compared to human recognition," personal communication, Oct. 2013.

[32] C. Landone, J. Harrop, and J. Reiss, "Enabling access to sound archives through integration, enrichment and retrieval: The Easaier project," in *Proc. 8th Int. Conf. Music Information Retrieval*, Vienna, Austria, Sept. 2007.

[33] L. Lu, H.-J. Zhang, and H. Jiang, "Content analysis for audio classification and segmentation," *IEEE Trans. Audio, Speech and Language Processing*, vol. 10, no. 7, pp. 504–516, 2002.

[34] R. G. Malkin and A. Waibel, "Classifying user environment for mobile applications using linear autoencoding of ambient audio," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (*ICASSP*), 2005, vol. 5, pp. 509–512.

[35] S. McAdams, "Recognition of sound sources and events," in *Thinking in Sound: The Cognitive Psychology of Human Audition,* S. McAdams and E. Bigand, Eds. London, U.K.: Oxford Univ. Press, 1993, pp. 146—98.

[36] Music information retrieval evaluation exchange (MIREX).[Online]. Available: http://music-ir.org/mirexwiki/

[37] V. T. Peltonen, A. J. Eronen, M. P. Parviainen, and A. P. Klapuri, "Recognition of everyday auditory scenes: Potentials, latencies and cues," in *Proc. 110th Audio Engineering Society Convention*, 2001, number 5404.

[38] V. T. Peltonen, J. T. Tuomi, A. Klapuri, J. Huopaniemi, and L. Sorsa, "Computational auditory scene recognition," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* (*ICASSP*), 2002, vol. 2, pp. 1941–1944.

[39] L. Rabiner and B.-H. Juang, *Fundamentals of Speech Recognition*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[40] R. Radhakrishnan, A. Divakaran, and P. Smaragdis, "Audio analysis for surveillance applications," in *Proc. IEEE Workshop Applications of Signal Processing to Audio and Acoustics* (*WASPAA*), 2005, pp. 158–161.

[41] M. Radovanović, A. Nanopoulos, and M. Ivanović, "Hubs in space: Popular nearest neighbors in high-dimensional data," *J. Mach. Learn. Res.*, vol. 11, pp. 2487–2531, Sept. 2010.

[42] N. Sawhney and P. Maes, "Situational awareness from environmental sounds," Techn. Rep., Massachusetts Institute of Technology, 1997.

[43] M. Schafer, *The Tuning of the World*. Rochester VT: Random House, 1977.

[44] R. E. Schapire, *The Boosting Approach to Machine Learning an Overview* (Lecture Notes in Statistics, vol. 171). New York: Springer, 2003, pp. 149–172.

[45] B. Schilit, N. Adams, and R. Want, "Context-aware computing applications," in *Proc. Workshop on Mobile Computing Systems and Applications*, 1994, pp. 85–90.

[46] (2013). Signal separation evaluation campaign (SiSEC). [Online]. Available: http://sisec.wiki.irisa.fr/tiki-index.php

[47] J. Tardieu, P. Susini, F. Poisson, P. Lazareff, and S. McAdams, "Perceptual study of soundscapes in train stations," *Appl. Acoust.*, vol. 69, no. 12, pp. 1224–1239, Dec. 2008.

[48] J. Thiemann, N. Ito, and E. Vincent, "The Diverse Environments Multichannel Acoustic Noise Database (DEMAND): A database of multichannel environmental noise recordings," in *Proc. 21st Int. Congr. on Acoustics*, Montreal, Canada, Acoustical Society of America, June 2013.

[49] M. van Grootel, T. Andringa, and J. Krijnders, "DARES-G1: Database of annotated real-world everyday sounds," in *Proc. NAG/DAGA Int. Conf. Acoustics*, 2009.

[50] P. Vandewalle, J. Kovacevic, and M. Vetterli, "Reproducible research in signal processing," *IEEE Signal Processing. Mag.*, vol. 26, no. 3, pp. 37–47, 2009.

[51] D. Wang and G. J. Brown, *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*. Piscataway, NJ: IEEE Press, 2006.

[52] Y. Xu, W. J. Li, and K. K. Lee, *Intelligent Wearable Interfaces*. Hoboken, NJ: Wiley, 2008.

[53] T. Zhang and C.-C. Jay Kuo, "Audio content analysis for online audiovisual data segmantation and classification," *IEEE Trans. Audio, Speech Lang. Processing*, vol. 9, no. 4, pp. 441–457, 2001.

[54] [Online]. Available: http://www.sound-ideas.com/sound-effects/series-6000-sound-effects-library.html

[55] D. Giannoulis, E. Benetos, D. Stowell, M. Rossignol, M. Lagrange, and M. D. Plumbley. (2013). IEEE AASP challenge: Detection and classification of acoustic scenes and events. [Online]. Available: http://c4dm.eecs.qmul.ac.uk/sceneseventschallenge/

[SP]

[Zhen-Hua Ling, Shi-Yin Kang, Heiga Zen, Andrew Senior, Mike Schuster,
Xiao-Jun Qian, Helen Meng, and Li Deng]

Deep Learning for Acoustic Modeling in Parametric Speech Generation

[A systematic review of existing techniques and future trends]

Hidden Markov models (HMMs) and Gaussian mixture models (GMMs) are the two most common types of acoustic models used in statistical parametric approaches for generating low-level speech waveforms from high-level symbolic inputs via intermediate acoustic feature sequences. However, these models have their limitations in representing complex, nonlinear relationships between the speech generation inputs and the acoustic features. Inspired by the intrinsically hierarchical process of human speech production and by the successful application of deep neural networks (DNNs) to automatic speech recognition (ASR), deep learning techniques have also been applied successfully to speech generation, as reported in recent literature. This article systematically reviews these emerging speech generation approaches, with the dual goal of helping readers gain a better understanding of the existing techniques as well as stimulating new work in the burgeoning area of deep learning for parametric speech generation.

In speech signal and information processing, many applications have been formulated as machine-learning tasks. ASR is a typical classification task that predicts word sequences from speech waveforms or feature sequences. There are also many regression tasks in speech processing that are aimed to generate speech signals from various types of inputs. They are referred to as *speech generation* tasks in this article. Speech generation covers a wide range of research topics in speech processing, such as text-to-speech (TTS) synthesis (generating speech from text), voice conversion (modifying nonlinguistic information of the input speech), speech enhancement (improving speech quality by noise reduction or other processing), and articulatory-to-acoustic mapping (converting articulatory movements to acoustic features). These

© ISTOCKPHOTO.COM/HUNG KUO CHUN

topics have the common goal of generating speech signals and differ in the forms of inputs. Statistical parametric speech generation (SPSG), which combines statistical acoustic models and vocoding techniques to generate speech waveforms, has been the mainstream approach for solving the speech generation problems. This approach first builds statistical acoustic models representing either the conditional probability density function (PDF) of output acoustic features given the input features or joint PDFs between the input and output features. The model structure is usually task dependent, but the parameters are estimated from a training database consisting of pairs of inputs and output acoustic features. At the speech-generation stage, the input features are given, which could be texts for TTS and noisy speech for speech enhancement. Then, the conditional distribution of the output acoustic features given the input features can be derived from the trained acoustic models. The output acoustic features are predicted from the conditional distribution under a certain criterion, e.g., maximizing the output probability, and are subsequently sent to a vocoder to reconstruct a speech waveform. In SPSG, vocoders are used to extract acoustic features, such as spectral [e.g., Mel-cepstral coefficients (MCCs)] and excitation (e.g., fundamental frequency and aperiodicity) features, from the raw waveforms of training data and to reconstruct speech waveforms from the generated acoustic features at synthesis time. Although both vocoder and acoustic modeling are essential for SPSG systems, this article focuses on acoustic modeling techniques for SPSG.

GMMs and HMMs with single Gaussian (or GMM) state-output PDFs are the two most popular acoustic models for SPSG [1], [2]. HMMs can represent nonstationary distributions of acoustic features using a sequence of hidden states, which are associated with linguistic features.

GMMs are widely used in frame-by-frame mapping for several speech-generation tasks, such as voice conversion, speech enhancement, and articulatory-to-acoustic mapping. The SPSG approaches using these two types of models have been shown to generate highly intelligible and smooth speech [2]–[4]. However, the generated speech sounds are noticeably muffled compared to recorded speech. Inadequate acoustic modeling is one of the main reasons for this deficiency [2], [5].

Take HMM-based speech synthesis, for example. In this approach, decision-tree-clustered, context-dependent phoneme HMMs are typically used to represent distributions of acoustic features given linguistic features [6]. The PDF of the acoustic features associated with each leaf node of the decision trees is typically a single Gaussian distribution with a diagonal covariance matrix.

At training time, parameters of the HMMs are usually estimated based on the maximum likelihood (ML) criterion. At synthesis time, given an input sentence and the trained parameters of the HMMs, the most likely acoustic features are predicted using the speech parameter-generation algorithm [7]. Since single Gaussian distributions are used as state-output PDFs, the outputs of the speech parameter-generation algorithm tend to distribute near the means of the Gaussian distributions, which are estimated by averaging all observations associated with a given decision tree leaf node. Although this averaging process improves the

robustness of parameter estimation and generation, the detailed characteristics of the speech parameters are often lost. Therefore, the reconstructed spectral envelopes are typically oversmoothed, which leads to the muffled voice quality of the synthetic speech. In recent years, many techniques have been proposed to alleviate the oversmoothing problem by introducing better acoustic models (e.g., the trajectory HMM [8], product of experts [9], and Gaussian process regression [10]), improving the model training criterion (e.g., minimum generation error training [11], [12]), or modifying the speech parameter-generation algorithm (e.g., integrating a global variance model [13], using segment-wise representation [14], and minimizing Kullback–Leibler divergences [15]).

Since 2006, deep learning has emerged as a new area of machine-learning research [16], [17] and has also attracted the attention of many signal processing researchers. Deep learning refers to a class of machine-learning techniques that exploit many layers of nonlinear information processing for supervised or unsupervised feature extraction and transformation, and for pattern analysis and classification. Both unconditional deep architectures [e.g., restricted Boltzmann machines (RBMs) [19], deep belief networks (DBNs) [16], denoising autoencoders (DAEs) [20], [21], deep Boltzmann machines [18], and conditional deep architectures, e.g., DNNs] [17], have been intensively studied and explored by signal processing researchers in recent years. Strictly speaking, an RBM is a shallow graphical model with only one layer of hidden units; it is the constituent of many deep models (e.g., DBNs and DNNs). As a density model, RBMs perform much better than the conventional shallow structures (e.g., GMMs) [18]. Considering its intrinsic relationship and similarity to other deep models, RBMs are included as an example of deep generative models in this article.

One example is the successful application of DNNs to the acoustic modeling of ASR. In this approach, DNNs are introduced to replace GMMs for evaluating the fit between a frame of acoustic observations and each HMM state [22]. Deep learning techniques have also been applied to the acoustic modeling of speech generation very recently to deal with the limitations of the conventional approaches [23]–[40]. Different from the deep learning in ASR where DNN-HMM is the dominant model structure, these emerging acoustic modeling approaches for speech generation adopted various model structures. Some of them focus on improving the density functions of HMM states or GMM mixtures using RBMs or DBNs [23], [24], [27]. While some others use DBNs or DNNs to model the entire mapping process from input to output feature sequences directly [25], [26], [28]–[35].

This article first reviews the conventional and popular statistical framework for speech generation, including HMM-based speech synthesis and GMM-based voice conversion, focusing on acoustic modeling and not on the vocoder. It then analyzes the limitations of these approaches. The key models and techniques of deep learning as relevant to speech generation, including RBMs, DBNs, and DNNs, are also introduced.

Subsequently, emerging speech generation approaches using deep learning techniques for acoustic modeling are reviewed systematically, with an analysis of their motivations and a

description of their implementations. Finally, we discuss the remaining issues associated with current deep learning methods for parametric speech generation and point to future directions in this area.

CONVENTIONAL ACOUSTIC MODELING USING HMMs AND GMMs FOR SPSG

HMM-BASED SPEECH SYNTHESIS

Statistical parametric speech synthesis (SPSS) [5] emerged in the mid-1990s [6], [41]. In this approach, the relationship between text and its acoustic realizations is modeled using a set of stochastic generative acoustic models. Decision-tree-clustered, context-dependent phoneme HMMs with single Gaussian state-output PDFs are the most popular generative acoustic model used in SPSS [6]. This approach is known as *HMM-based speech synthesis*. An HMM is a generative model that generates an observation sequence using a discrete and hidden state sequence. An example of a three-state left-to-right HMM is illustrated in Figure 1. In an HMM, state-output PDFs describe the distribution of observed features belonging to corresponding states and the transition among states is characterized by state-transition probabilities.

HMM-based speech synthesis is able to synthesize highly intelligible and smooth speech sounds. In addition, this model-based approach makes speech synthesis far more flexible compared to the conventional unit selection and waveform concatenation approach. Model adaptation, interpolation, and manipulation methods have been applied to control the HMM's
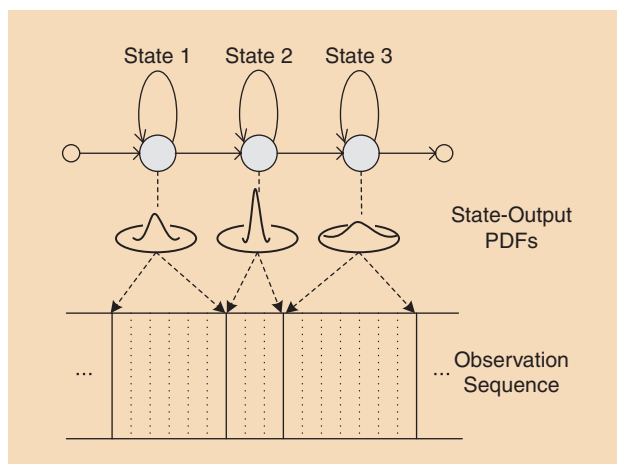
[FIG1] An example of a three-state, left-to-right HMM.

parameters and thus diversify the characteristics of the generated speech [42]–[49]. Figure 2 shows the diagram of a typical HMM-based speech synthesis system. At the training stage, acoustic features of speech, including vocal tract and vocal source parameters, are extracted from the speech waveforms in a training database. Context features are also derived from the segmental and prosodic labels of the texts corresponding to the waveforms. Then, a set of parameters of context-dependent HMMs λ^* is estimated based on the ML criterion as

$$\lambda^* = \arg \max_{\lambda} p(y \mid x, \lambda), \tag{1}$$
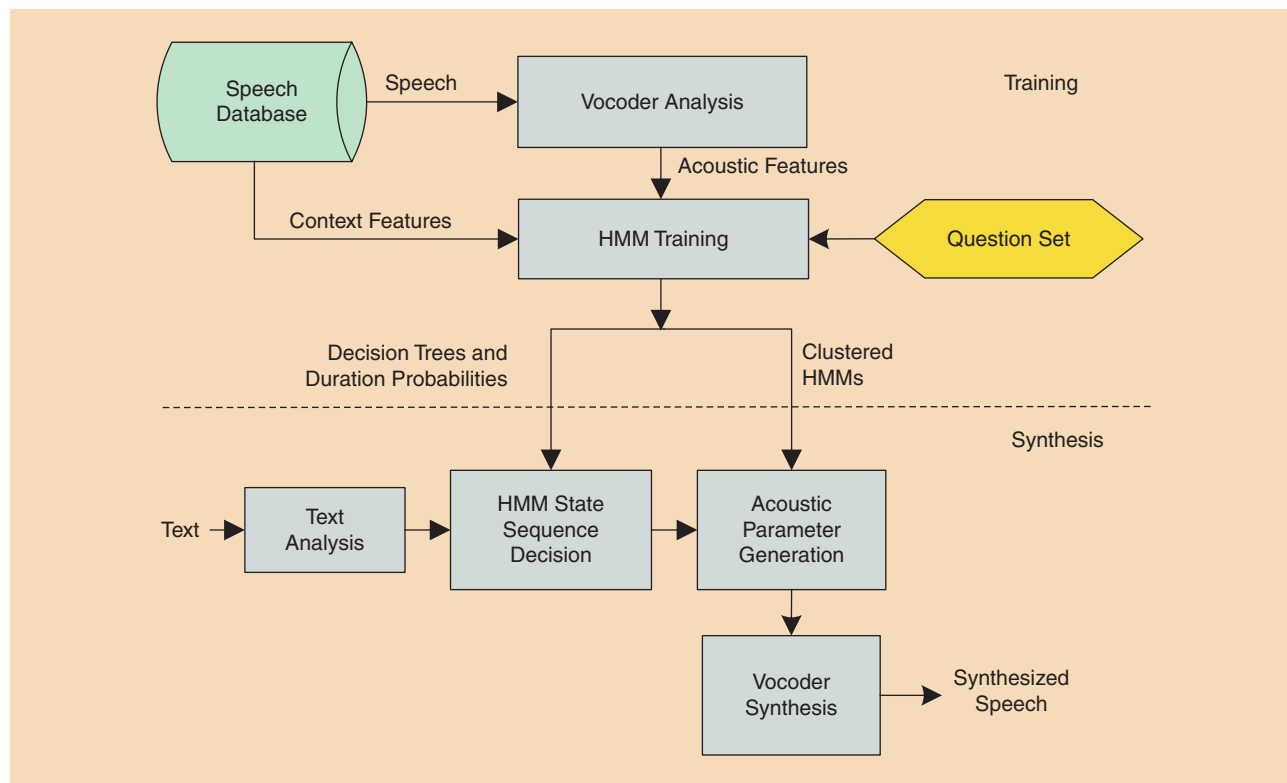
[FIG2] A block diagram of a typical HMM-based speech synthesis system.

where $p(\cdot)$ is used to denote a PDF (continuous) in this article, $\boldsymbol{y} = [\boldsymbol{y}_1^\top, \boldsymbol{y}_2^\top, \ldots, \boldsymbol{y}_T^\top]^\top$ denotes a sequence of acoustic features with T frames, $\boldsymbol{y}_t$ is the acoustic feature at frame t, $\boldsymbol{x} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ is a sequence of linguistic context features for $\boldsymbol{y}$ that are derived from text automatically or annotated manually, N is the number of phonemes, and $(\cdot)^\top$ denotes the matrix transposition operation. The acoustic feature vector at each frame typically consists of static acoustic parameters $\boldsymbol{y}_{s_t} \in \mathcal{R}^{D_y}$ and their velocity and acceleration components, $\Delta \boldsymbol{y}_{s_t}$ and $\Delta^2 \boldsymbol{y}_{s_t}$, as

$$\boldsymbol{y}_t = [\boldsymbol{y}_{s_t}^\top, \Delta \boldsymbol{y}_{s_t}^\top, \Delta^2 \boldsymbol{y}_{s_t}^\top]^\top. \tag{2}$$

Therefore, the complete acoustic feature sequence $\boldsymbol{y}$ can be considered a linear transform of the static feature sequence $\boldsymbol{y}_s = [\boldsymbol{y}_{s_1}^\top, \boldsymbol{y}_{s_2}^\top, \ldots, \boldsymbol{y}_{s_T}^\top]^\top$ as

$$\boldsymbol{y} = M_y \boldsymbol{y}_s, \tag{3}$$

where M_y is determined by the velocity and acceleration calculation functions used in (2) [7].

An HMM-based speech synthesis system typically contains a large number of context-dependent HMMs with linguistic context features that are far more extensive and can express far more fine-grained distinctions than those used in HMM-based ASR systems [50], [51]. This leads to data sparsity problems, such as overfitting in context-dependent models that have only few training examples available and the problem that many valid combinations of linguistic context features will be absent from the training database. To deal with this issue, a decision-tree-based clustering technique [52] is applied after the initial training to cluster state-output PDFs of the context-dependent HMMs as shown in Figure 3, where

[FIG3] A decision-tree-based modeling clustering for HMM-based speech synthesis.

the state-output PDFs of the context-dependent HMMs with similar context descriptions are represented by a shared distribution. The question set for decision tree constriction is designed considering the characteristics of the language being processed. Next, the state alignment results using the trained HMMs are utilized to train context-dependent state-duration PDFs [6]. A single Gaussian distribution is also used to model the state-duration PDF at each state. A decision-tree-based model clustering technique is similarly applied to these state-duration PDFs [54]. Joint training of state-output and state-duration PDFs based on hidden semi-Markov models have also been used [53].

The acoustic model $p(\boldsymbol{y} \mid \boldsymbol{x}, \lambda)$ used in HMM-based speech synthesis can be rewritten as

$$p(\boldsymbol{y} \mid \boldsymbol{x}, \lambda) = \sum_{\forall \boldsymbol{q}} p(\boldsymbol{y}, \boldsymbol{q} \mid \boldsymbol{x}, \lambda), \tag{4}$$

$$= \sum_{\forall \boldsymbol{q}} P(\boldsymbol{q} \mid \boldsymbol{x}, \lambda) p(\boldsymbol{y} \mid \boldsymbol{q}, \lambda), \tag{5}$$

$$= \sum_{\forall \boldsymbol{q}} P(\boldsymbol{q} \mid \boldsymbol{x}, \lambda) \prod_{t=1}^{T} p(\boldsymbol{y}_t \mid \boldsymbol{q}_t, \lambda), \tag{6}$$

where $P(\cdot)$ is used to denote a probability mass function (discrete) in this article, $p(\boldsymbol{y}_t \mid \boldsymbol{q}_t, \lambda)$ is a state-output PDF associated with the q_tth state, which is typically a single Gaussian distribution with a diagonal covariance matrix and $\boldsymbol{q} = \{q_1, \ldots, q_T\}$ is an HMM state sequence. Note that the derivation from (5) to (6) is based on the assumption of HMMs that the frame observations are independent from each other given the state sequence.

To perform synthesis, the result of front-end linguistic analysis on input text is used to get the context features $\tilde{\boldsymbol{x}}$ for synthesis, as shown in Figure 2. In the HMM state sequence decision step, a sentence HMM corresponding to the input text is composed, with its parameters derived from the training stage.

In the step of acoustic parameter-generation, the acoustic features that maximize their output probabilities given the sentence HMM are determined under the constraints between static and dynamic features [7] as

$$\boldsymbol{y}_s^* = \arg \max_{\boldsymbol{y}_s} p(\boldsymbol{y} \mid \tilde{\boldsymbol{x}}, \lambda^*) \Big|_{\boldsymbol{y} = M_y \boldsymbol{y}_s}. \tag{7}$$

The solution to (7) can be simplified if only the optimal state sequences in (5) is considered; optimization is approximated as two sequential steps

$$\boldsymbol{q}^* = \arg \max_{\boldsymbol{q}} P(\boldsymbol{q} \mid \tilde{\boldsymbol{x}}, \lambda^*), \tag{8}$$

$$\boldsymbol{y}_s^* = \arg \max_{\boldsymbol{y}_s} \prod_{t=1}^{T} p(\boldsymbol{y}_t \mid \boldsymbol{q}_t^*, \lambda^*) \Big|_{\boldsymbol{y} = M_y \boldsymbol{y}_s}. \tag{9}$$

Then, the closed-form solution of $\boldsymbol{y}_s^*$ can be derived by setting the partial derivative of (9) with respect to $\boldsymbol{y}_s$ to **zero** once the state sequence $\boldsymbol{q}^*$ is given [7]. Finally, these generated parameters are sent to a vocoder to reconstruct the speech waveforms.

GMM-BASED VOICE CONVERSION
The aim of voice conversion is to modify the nonlinguistic information (e.g., speaker characteristics) of input speech while
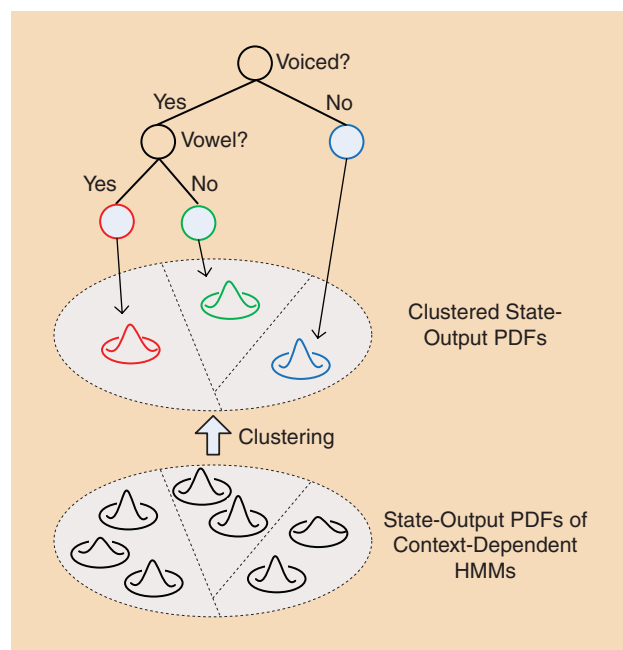
keeping the linguistic information unchanged. Different from the linguistic features, which are used as inputs for speech synthesis, the input features for voice conversion are typically continuous acoustic representations of a source voice. Many statistical approaches to voice conversion have been studied since the late 1980s, such as codebook mapping [55], GMM [2], [56], frequency warping [57], neural networks [58], partial least square regression [59], noisy channel model [60], etc. Among them, GMM-based voice conversion is the most popular [2], [56]. Figure 4 is a diagram of a typical GMM-based voice conversion system with parallel training data, which means that the training database contains the speech waveforms uttered by the source and target voices for the same texts. At the training stage, the acoustic features of the source and target speech in the training database are extracted by a vocoder and are aligned frame by frame by dynamic time warping. Then, the aligned pairs of the source acoustic feature vector x_t and the target acoustic feature vector y_t are concatenated to construct a joint feature vector $z_t = [x_t^\top, y_t^\top]^\top$. Similar to HMM-based speech synthesis, the acoustic features x_t and y_t consist of static and dynamic components. Therefore, the acoustic feature sequences $x = [x_1^\top, x_2^\top, ..., x_T^\top]^\top$ and $y = [y_1^\top, y_2^\top, ..., y_T^\top]^\top$ can also be written as a linear transform from the static feature sequences $x_s = [x_{s_1}^\top, x_{s_2}^\top, ..., x_{s_T}^\top]^\top$ and $y_s = [y_{s_1}^\top, y_{s_2}^\top, ..., y_{s_T}^\top]^\top$ as $x = M_x x_s$ and $y = M_y y_s$, where M_x and M_y are determined by the velocity and acceleration calculation functions [2]. Then, a joint distribution GMM (JD-GMM) λ with a set of parameters $\{\alpha_m, \mu_m^{(z)}, \Sigma_m^{(z)}\}_{m=1}^M$ is estimated to model

a joint PDF between the source and target acoustic features, where M denotes the total number of mixture components in the JD-GMM, and α_m, $\mu_m^{(z)}$, and $\Sigma_m^{(z)}$ correspond to the mixture weight, mean vector, and covariance matrix associated with the mth Gaussian component. The mean vector and covariance matrix are structured as

$$\mu_m^{(z)} = \begin{bmatrix} \mu_m^{(x)} \\ \mu_m^{(y)} \end{bmatrix}, \Sigma_m^{(z)} = \begin{bmatrix} \Sigma_m^{(xx)} & \Sigma_m^{(xy)} \\ \Sigma_m^{(yx)} & \Sigma_m^{(yy)} \end{bmatrix}. \tag{10}$$

To reduce the number of model parameters and computational cost, $\Sigma_m^{(xx)}$, $\Sigma_m^{(yy)}$, $\Sigma_m^{(xy)}$, and $\Sigma_m^{(yx)}$ are commonly set to be diagonal [2]. These model parameters are typically estimated by the ML criterion as

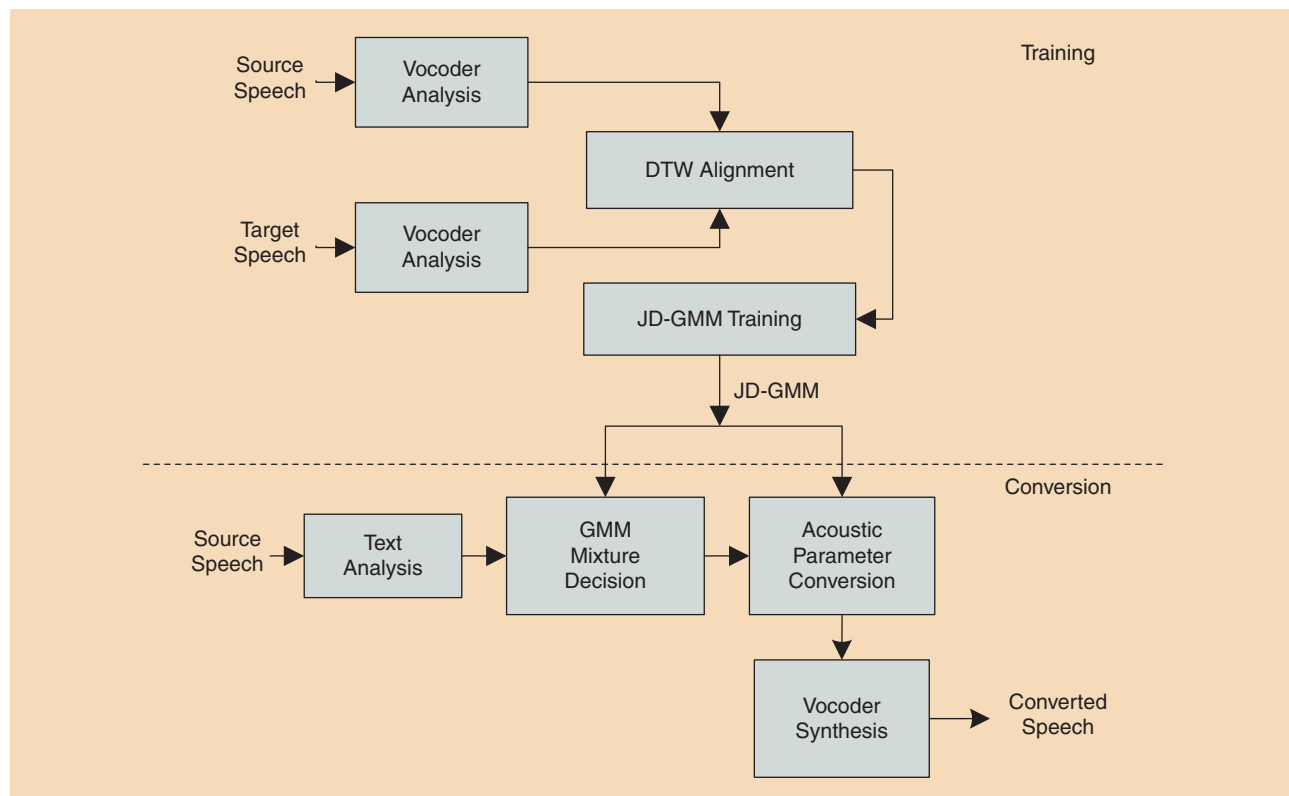$$\lambda^* = \arg \max_\lambda p(x, y \mid \lambda), \tag{11}$$

$$= \arg \max_\lambda \prod_{t=1}^T p(z_t \mid \lambda). \tag{12}$$

The conditional PDF given an input source acoustic feature $\tilde{x}$ can be further derived from the trained JD-GMM λ^* as

$$p(y \mid \tilde{x}, \lambda^*) = \sum_{\forall m} p(y, m \mid \tilde{x}, \lambda^*), \tag{13}$$

$$= \sum_{\forall m} P(m \mid \tilde{x}, \lambda^*) \prod_{t=1}^T p(y_t \mid \tilde{x}_t, m_t, \lambda^*), \tag{14}$$

where $m = \{m_1, ..., m_T\}$ denotes the sequence of mixture components. $P(m \mid \tilde{x}, \lambda^*) = \prod_{t=1}^T P(m_t \mid \tilde{x}_t, \lambda^*)$ and $P(m_t \mid \tilde{x}_t, \lambda^*)$

[FIG4] A block diagram of a typical GMM-based voice conversion system.

can be determined from the marginal PDF of x_t, which is a GMM of M mixture components with the set of model parameters $\{\alpha_m, \boldsymbol{\mu}_m^{(x)}, \Sigma_m^{(x)}\}$. The conditional PDF $p(y_t | x_t, m_t, \lambda)$ is a Gaussian distribution with a mean vector

$$\boldsymbol{\mu}_{m,t}^{y|x} = \boldsymbol{\mu}_m^{(y)} + \Sigma_m^{(yx)} \Sigma_m^{(xx)-1} (x_t - \boldsymbol{\mu}_m^{(x)}) \qquad (15)$$

and a covariance matrix

$$\Sigma_m^{y|x} = \Sigma_m^{(yy)} - \Sigma_m^{(yx)} \Sigma_m^{(xx)-1} \Sigma_m^{(xy)}. \qquad (16)$$

Figure 5(a) shows the PDF of an example JD-GMM with two mixtures, where the source and target acoustic features are simply represented by scalars. Two examples of the conditional distributions derived from the JD-GMM are illustrated in Figure 5(b), which are also two-mixture GMMs.

At conversion time, the converted acoustic features can be predicted using either the minimum mean-square error [56] or the maximum a posteriori criterion [2], given the source acoustic feature sequence $\tilde{x}$. If the maximum a posteriori criterion is adopted, the static acoustic features of the target voice are predicted as

$$y_s^* = \arg\max_{y_s} p(y | \tilde{x}, \lambda^*)\Big|_{y = M_y y_s}. \qquad (17)$$

Similar to HMM-based speech synthesis, the solution to (17) is simplified by only considering the mixture components with the highest posterior probability at each frame in (14). Thus, we have

$$m_t^* = \arg\max_{m_t} P(m_t | \tilde{x}_t, \lambda^*), \qquad (18)$$

$$y_s^* = \arg\max_{y_s} \prod_{t=1}^{T} p(y_t | \tilde{x}_t, m_t^*, \lambda^*)\Big|_{y = M_y y_s}. \qquad (19)$$
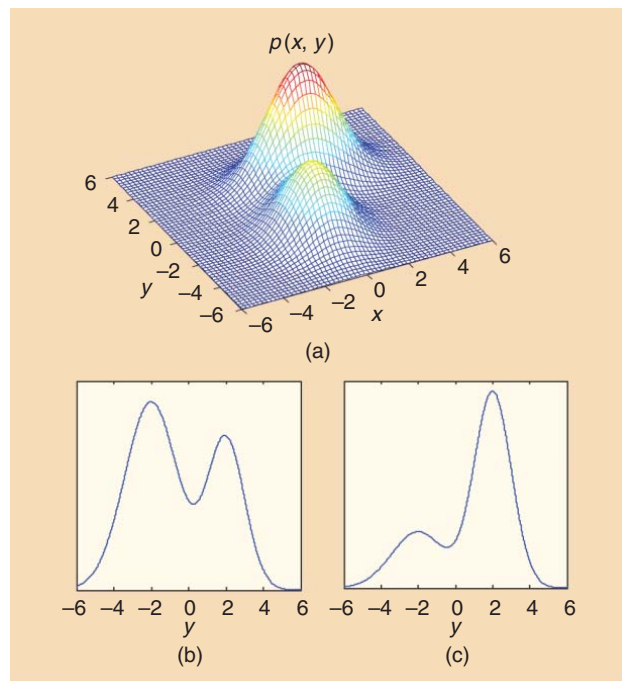
[FIG5] PDFs of (a) a joint distribution GMM $p(x, y)$ with two mixtures and (b) and (c) the conditional distributions $p(y | x)$ derived from it. (b) $p(y | x = -1)$. (c) $p(y | x = 1)$.

Then, a closed-form solution to (19) can be achieved in a similar way to solve (9) [2]. Finally, the converted acoustic features are sent to a vocoder to reconstruct the corresponding speech waveform.

This GMM-based voice conversion framework has also been successfully applied to other frame-by-frame-mapping speech generation tasks, such as bandwidth extension [61], speech enhancement [62], [63], and articulatory-acoustic mapping [64].

THE COMMON STRUCTURE: TWO-STEP MAPPING

As shown in (8), (9), (18), and (19), both HMM- and GMM-based SPSG share the common structure of two-step mapping to represent the conditional PDF of the acoustic features y, given the input features x.

1) *Input-to-cluster mapping using hidden discrete variable*: In this step, each input feature vector is mapped to hidden discrete clusters of the acoustic features to be generated, i.e., the HMM state q_t^* in (8) or the GMM mixture component m_t^* in (18). In HMM-based speech synthesis, q^* is determined using the decision trees for state-output PDFs and the state-duration PDFs. In GMM-based voice conversion, this is achieved by the posterior probabilities $P(m_t | \tilde{x}_t, \lambda^*)$.

2) *Cluster-to-feature mapping using Gaussian distributions*: Given the input features, once the cluster sequence is determined, the conditional PDF for generating the acoustic features can be determined by combining the PDFs describing each cluster in the sequence, i.e., $p(y_t | q_t^*, \lambda^*)$ in (9) and $p(y_t | \tilde{x}_t, m_t^*, \lambda^*)$ in (19). In the current SPSG approaches, the PDF associated with each cluster is typically an ML-estimated single Gaussian distribution with a diagonal covariance matrix [2], [6].

Although the acoustic modeling approach described earlier works reasonably well in SPSG, it has well known limitations. First, decision-tree-based input-to-cluster mapping in HMM-based speech synthesis is inefficient for expressing complex context dependencies, such as the exclusive OR (XOR) problem. This may lead to overfitting to the training data because of the data partitioning issue [65]. Second, the cluster-to-feature mapping using single Gaussian distributions with diagonal covariance matrices is established based on two independence assumptions: 1) conditional independence between frames given the state or the Gaussian component and 2) independence of acoustic features within a frame. As discussed earlier, this leads to reconstructed spectral envelopes being oversmoothed and the quality of synthetic speech is degraded.

Compared with the statistical models used in the conventional acoustic modeling of SPSG (such as decision trees, HMMs, and GMMs), deep learning techniques are better at representing the intrinsic correlations among the units of input vectors (e.g., the input context features for speech synthesis), among the units of output vectors (e.g., the output spectral features for speech synthesis), and between the input and output vectors (e.g., the aligned spectral features of the source and target speakers for voice conversion) using a joint (e.g., RBM and DBN) or conditional (e.g., DNN) modeling framework. Therefore, it is promising that the deep learning techniques can help the acoustic modeling of speech generation to

overcome the limitations of the current approach mentioned earlier, so as to achieve better input-to-cluster or/and cluster-to-feature mapping. Furthermore, human speech production mechanisms involve clearly layered hierarchical structures in transforming the information from the linguistic level to the acoustic level via intermediate levels of motor control and articulation [66]–[69], also suggesting the need for deep model structures for SPSG applications.

This article reviews a number of recent approaches, based on the deep learning techniques, for overcoming these limitations and improving acoustic modeling for SPSG. A few basic models for deep learning are first reviewed in the section "Basic Models for Deep Learning," including some mathematical details that are uncommon in the literature but essential for using these models in SPSG.

BASIC MODELS FOR DEEP LEARNING

Since 2010, deep learning techniques have been successfully applied to the modeling of speech signals, such as speech recognition [70]–[74], spectrogram coding [20], voice activity detection [75], and acoustic-articulatory inversion mapping [76]. One significant advantage of deep learning techniques is their strong ability to represent the intrinsic correlation or mapping relationship among the units of a high-dimensional stochastic vector using a joint (e.g., RBM and DBN) or conditional (e.g., CRBM and DNN) modeling framework. Considering that speech generation is a regression task and the aim of its acoustic modeling is to describe the joint or conditional distribution of continuous acoustic features, we will review these basic models from the viewpoint of density models in this section.

RBMs

An RBM is an undirected graphical model (i.e., a Markov random field) that can model the dependency among a set of random variables using a two-layered architecture [19]. In an RBM, visible stochastic units $v = [v_1, \ldots, v_V]^\top$ are connected to hidden stochastic units $h = [h_1, \ldots, h_H]^\top$, as shown in Figure 6, where V and H are the numbers of units at the visible and hidden layers, respectively. When $v \in \{0,1\}^V$ and $h \in \{0,1\}^H$ are both binary stochastic variables, the energy function of the state $\{v, h\}$ is defined as

$$E(v, h; \lambda) = -\sum_{i=1}^{V} a_i v_i - \sum_{j=1}^{H} b_j h_j - \sum_{i=1}^{V} \sum_{j=1}^{H} w_{ij} v_i h_j, \quad (20)$$

where w_{ij} represents the symmetric interaction between v_i and h_j, a_i and b_j are bias terms, and λ denotes the set of model parameters consisting of $a = [a_1, \ldots, a_V]^\top$, $b = [b_1, \ldots, b_H]^\top$, and $W = \{w_{ij}\} \in \mathcal{R}^{V \times H}$. The joint PDF over the visible and hidden units is given by a Boltzmann distribution as

$$P(v, h \mid \lambda) = \frac{1}{\mathcal{Z}_\lambda} \exp\{-E(v, h; \lambda)/C_T\}, \quad (21)$$

where C_T is a temperature parameter, which is assumed to be 1 in the rest of this article, and

$$\mathcal{Z}_\lambda = \sum_{\forall v} \sum_{\forall h} \exp\{-E(v, h; \lambda)\} \quad (22)$$
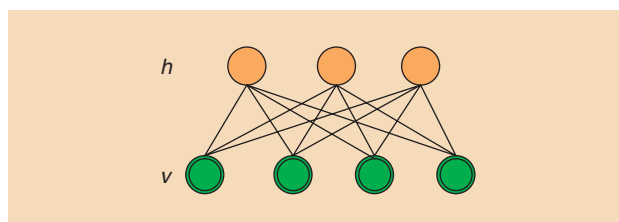
[FIG6] A graphical model representation for an RBM.

is the partition function, which can be estimated using the annealed importance sampling (AIS) technique [18]. The marginal PDF over the visible vector v can be calculated as

$$P(v \mid \lambda) = \frac{1}{\mathcal{Z}_\lambda} \sum_{\forall h} \exp\{-E(v, h; \lambda)\}. \quad (23)$$

Given a training set, λ can be estimated based on the ML criterion by stochastic gradient descent. The derivative of $\log P(v \mid \lambda)$ with respect to the model parameters, e.g., w_{ij}, can be derived using (20)–(23) as

$$\frac{\partial \log P(v \mid \lambda)}{\partial w_{ij}} = \mathrm{E}_{P_{\mathrm{Data}}}[v_i h_j] - \mathrm{E}_{P_{\mathrm{Model}}}[v_i h_j], \quad (24)$$

where $\mathrm{E}_{P_{\mathrm{Data}}}[\cdot]$ denotes an expectation with respect to the distribution of the training data and $\mathrm{E}_{P_{\mathrm{Model}}}[\cdot]$ denotes an expectation with respect to the distribution of the model $P(v \mid \lambda)$. Because computation of $\mathrm{E}_{P_{\mathrm{Model}}}[\cdot]$ is intractable, the contrastive divergence (CD) algorithm has been proposed to approximate $\mathrm{E}_{P_{\mathrm{Model}}}[\cdot]$ by Gibbs sampling [77].

RBMs can also be applied to model the distribution of real-valued data (e.g., mel-frequency MCCs in ASR), categorical data (e.g., some linguistic context features in TTS), or a mixed vector of binary, real-valued, and categorical data by defining different forms of energy functions [25]. For a Gaussian-Bernoulli RBM, which means $v \in \mathcal{R}^V$ are real-valued and $h \in \{0,1\}^H$ are binary, the energy is defined as

$$E(v, h; \lambda) = \sum_{i=1}^{V} \frac{(v_i - a_i)^2}{2\sigma_i^2} - \sum_{j=1}^{H} b_j h_j - \sum_{i=1}^{V} \sum_{j=1}^{H} w_{ij} h_j \frac{v_i}{\sigma_i}, \quad (25)$$

where the variance parameters σ_i^2 are commonly fixed to a predetermined value instead of learning them from training data [17]. While training a Gaussian–Bernoulli RBM using the CD algorithm, the two conditional PDFs for Gibbs sampling are derived as

$$P(h_j = 1 \mid v, \lambda) = g\left(b_j + v^\top \Sigma^{-\frac{1}{2}} w_{\cdot j}\right), \quad (26)$$
$$p(v \mid h, \lambda) = \mathcal{N}(v; Wh + a, \Sigma), \quad (27)$$

where $g(x) = 1/(1 + \exp(-x))$ is a sigmoid function, $w_{\cdot j}$ denotes the jth column of a matrix W, $\mathcal{N}(v; \mu, \Sigma)$ denotes a Gaussian distribution of v with a mean vector μ and a covariance matrix Σ, and $\Sigma = \mathrm{diag}\{\sigma_1^2, \ldots, \sigma_V^2\}$ is diagonal. If $\{\sigma_i^2\}_{i=1}^{V}$ are fixed to 1, Σ turns into an identity matrix.

RBMs have been successfully used in unsupervised pretraining of DNN-based acoustic models in ASR [22]. RBMs have also been used as density models to represent the distributions of acoustic

features for SPSG [23], [24], [27]. The marginal PDF of a Gaussian–Bernoulli RBM can be derived from (23) and (25) as [the variance parameters σ_i^2 in (25) are fixed to 1 for notational simplicity]

$$
\begin{aligned}
p(v \mid \lambda) &= \frac{1}{\mathcal{Z}_\lambda} \sum_{\forall h} \exp\{-E(v, h; \lambda)\} \\
&= \frac{1}{\mathcal{Z}_\lambda} \sum_{\forall h} \exp\left\{-\sum_{i=1}^{V} \frac{(v_i - a_i)^2}{2} + b^\top h + v^\top W h\right\} \\
&= \frac{1}{\mathcal{Z}_\lambda} \exp\left\{-\sum_{i=1}^{V} \frac{(v_i - a_i)^2}{2}\right\} \\
&\quad \cdot \prod_{j=1}^{H} \sum_{h_j \in \{0,1\}} \exp(b_j h_j + v^\top w_{\cdot j} h_j) \\
&= \frac{1}{\mathcal{Z}_\lambda} \prod_{i=1}^{V} \exp\left\{-\frac{(v_i - a_i)^2}{2}\right\} \\
&\quad \cdot \prod_{j=1}^{H} \{1 + \exp(b_j + v^\top w_{\cdot j})\},
\end{aligned}
\tag{28}
$$

which shows that a Gaussian–Bernoulli RBM can be considered either a product of experts (PoEs) or a GMM.

■ *PoE* [78]: A PoE represents a probability distribution by multiplying several simpler distributions, followed by normalization. PoEs can produce much sharper distributions than their individual experts and perform more efficiently than mixture models in high-dimensional space [77]. As shown in (28), elements in the first product represent single-variable experts without cross-dimensional correlations. The elements in the second product represent constraints between input variable using the model parameters corresponding to each hidden unit.

■ *GMM*: An RBM can also be considered as a GMM with 2^H mixture components with structured mean vectors and identity covariance matrices. For example, if $H = 0$,

$$
p(v \mid \lambda) = \frac{1}{\mathcal{Z}_\lambda} \exp\left\{-\sum_{i=1}^{V} \frac{(v_i - a_i)^2}{2}\right\}
\tag{29}
$$

is a single Gaussian distribution with a mean vector a. If H is increased to 1, $p(v \mid \lambda)$ in (28) can be rewritten as

$$
\begin{aligned}
p(v \mid \lambda) &= \frac{1}{\mathcal{Z}_\lambda} \exp\left\{-\sum_{i=1}^{V} \frac{(v_i - a_i)^2}{2}\right\} \\
&\quad + \frac{\kappa}{\mathcal{Z}_\lambda} \exp\left\{-\sum_{i=1}^{V} \frac{(v_i - a_i - w_{i1})^2}{2}\right\},
\end{aligned}
\tag{30}
$$

where κ is a constant value determined by the model parameters. We can see that $p(v \mid \lambda)$ becomes a GMM with two mixture
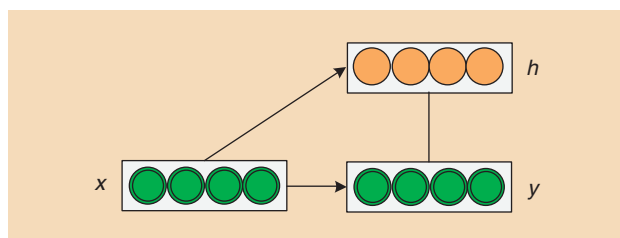
[FIG7] The graphical model representation for a CRBM.

components, where their mean vectors become a and $a + w_{\cdot 1}$, respectively. Generally speaking, as the number of hidden units is incremented, the number of mixture components is doubled by copying and shifting the mean vectors. These structured mean vectors and the tied covariance matrices provide better generalization. Thus, they are robust toward data sparsity.

RBMs can also be used to model conditional PDFs between two groups of visible units using their variation form, i.e., the conditional RBM (CRBM). The CRBM was originally proposed to model the temporal dependency of human motion features [79]. The model structure of a CRBM representing the conditional PDF $p(y \mid x, \lambda)$ is illustrated in Figure 7. In this model, the links between the visible units y and the hidden units h are undirected. If x is known, y and h form an RBM and its model parameters depend on x through the two directed links from x to y and h. If $h \in \{0,1\}^H$ are binary and $x \in \mathcal{R}^{D_X}$ and $y \in \mathcal{R}^{D_Y}$ are real-valued, the energy function of a CRBM can be written as

$$
\begin{aligned}
E(y, h, x; \lambda) &= \sum_{i=1}^{D_Y} \frac{\left(y_i - a_i - \sum_k A_{ki} x_k\right)^2}{2\sigma_i^2} \\
&\quad - \sum_{j=1}^{H} \left(b_j + \sum_k B_{kj} x_k\right) h_j - \sum_{i=1}^{D_Y} \sum_{j=1}^{H} w_{ij} h_j \frac{y_i}{\sigma_i},
\end{aligned}
\tag{31}
$$

where $\lambda = \{A, B\}$ is the set of parameters in the CRBM, $A = \{A_{ki}\} \in \mathcal{R}^{D_X \times V}$ and $B = \{B_{kj}\} \in \mathcal{R}^{D_X \times H}$ are matrices corresponding to the directed links in Figure 7. The conditional PDF of y given x can be written as

$$
p(y \mid x, \lambda) = \sum_{\forall h} p(y, h \mid x, \lambda)
\tag{32}
$$

$$
= \frac{1}{\mathcal{Z}_\lambda} \sum_{\forall h} \exp\{-E(y, h, x; \lambda)\},
\tag{33}
$$

where

$$
p(y, h \mid x, \lambda) = \frac{1}{\mathcal{Z}_\lambda} \exp\{-E(y, h, x; \lambda)\},
\tag{34}
$$

$$
\mathcal{Z}_\lambda = \int \sum_{\forall h} \exp\{-E(y, h, x; \lambda)\} dy.
\tag{35}
$$

Similar to RBMs, λ can be trained based on the ML criterion using the CD algorithm [79].

DBNs
A DBN is a probabilistic generative model that is composed of many layers of hidden units [16]. The graphical model representation for a three-hidden-layer DBN is shown in Figure 8. In this model, each layer captures the correlations among the activities of hidden features in the layer below. The top two layers of the DBN form an undirected graph. The lower layers form a directed graph with a top–down direction to generate the visible units. Assuming that v is real-valued and $\{h^{(l)}\}_{l=1}^{L}$ are binary, the joint PDF of a DBN over the visible and hidden units can be written as

$$
\begin{aligned}
p(v, h^{(1)}, \ldots, h^{(L)} \mid \lambda) &= p(v \mid h^{(1)}, \lambda) \prod_{l=2}^{L-1} P(h^{(l-1)} \mid h^{(l)}, \lambda) \\
&\quad \cdot P(h^{(L-1)}, h^{(L)} \mid \lambda),
\end{aligned}
\tag{36}
$$

where $\boldsymbol{h}^{(l)} = [h_1^{(l)}, \ldots, h_{H_l}^{(l)}]^\top$ is the hidden stochastic vector at the lth hidden layer, H_l is the dimensionality of $\boldsymbol{h}^{(l)}$, and L is the number of hidden layers. $P(\boldsymbol{h}^{(L-1)}, \boldsymbol{h}^{(L)} \mid \lambda)$ is represented by an RBM as (21) with the weight matrix $W^{(L)}$ and the bias vectors $\boldsymbol{a}^{(L)}$ and $\boldsymbol{b}^{(L)}$. $p(\boldsymbol{v} \mid \boldsymbol{h}^{(1)}, \lambda)$ and $\{P(\boldsymbol{h}^{(l-1)} \mid \boldsymbol{h}^{(l)}, \lambda)\}_{l=2}^{L-1}$ are represented by sigmoid belief networks [80]. Each sigmoid belief network is described by a weight matrix $W^{(l)}$ and a bias vector $\boldsymbol{a}^{(l)}$. Assuming that $\boldsymbol{v}$ is real-valued and $\{\boldsymbol{h}^{(l)}\}_{l=2}^{L}$ are binary, the conditional PDF $p(\boldsymbol{v} \mid \boldsymbol{h}^{(1)}, \lambda)$ of a sigmoid belief network is described by (27). For $l \in \{2, 3, \ldots, L-1\}$, the dependency between two adjacent hidden layers is represented by

$$P(h_i^{(l-1)} = 1 \mid \boldsymbol{h}^{(l)}, \lambda) = g\left(a_i^{(l)} + \sum_j w_{ij}^{(l)} h_j^{(l)}\right). \tag{37}$$

For an L-hidden-layer DBN, its model parameters are composed of $\{\boldsymbol{a}^{(1)}, W^{(1)}, \ldots, \boldsymbol{a}^{(L-1)}, W^{(L-1)}, \boldsymbol{a}^{(L)}, \boldsymbol{b}^{(L)}, W^{(L)}\}$. Furthermore, the marginal PDF of the visible variables for a DBN can be written as

$$p(\boldsymbol{v} \mid \lambda) = \sum_{\forall \boldsymbol{h}^{(l)}} \cdots \sum_{\forall \boldsymbol{h}^{(L)}} p(\boldsymbol{v}, \boldsymbol{h}^{(1)}, \ldots, \boldsymbol{h}^{(L)} \mid \lambda). \tag{38}$$

Given the training samples of the visible units, it is difficult to estimate the model parameters of a DBN directly based on the ML criterion due to the complex model structure with multiple hidden layers. Therefore, a greedy learning algorithm has been proposed and popularly applied to train DBNs in a layer-by-layer manner [16]. A stack of RBMs are used in this algorithm. First, it estimates the parameters $\{\boldsymbol{a}^{(1)}, \boldsymbol{b}^{(1)}, W^{(1)}\}$ of the first-layer RBM to model the visible training data. Then, it freezes the parameters $\{\boldsymbol{a}^{(1)}, W^{(1)}\}$ of the first layer and draws samples from $P(\boldsymbol{h}^{(1)} = 1 \mid \boldsymbol{v}, \lambda)$ using (26) to train the next-layer RBM $\{\boldsymbol{a}^{(2)}, \boldsymbol{b}^{(2)}, W^{(2)}\}$. This training procedure is conducted recursively until it reaches the top layer and gets $\{\boldsymbol{a}^{(L)}, \boldsymbol{b}^{(L)}, W^{(L)}\}$. It has been shown that this greedy learning algorithm can improve the lower bound on the log-likelihood of the model, given training samples by adding each new hidden layer [16], [18]. Once the model parameters are estimated, the calculation of the log probability that a DBN assigns to training or test data by applying (38) directly becomes computationally intractable. A lower bound on the log probability can be estimated by combining the AIS-based partition function estimation with the approximate inference [18].

DNNs

A DNN is a feed-forward, artificial neural network that has more than one layer of hidden units between its input and output layers [22]. The model representation for a two-hidden-layer DNN is shown in Figure 9. At each hidden layer, each hidden unit typically maps the weighted sum of its inputs from the layer below to a deterministic value using a nonlinear activation function and passes it to the layer above. If a sigmoid function $g(\cdot)$ is used as an activation function, its output is given as

$$h_j^{(l)} = g\left(b_j^{(l)} + \sum_i h_i^{(l-1)} w_{ij}^{(l)}\right), \tag{39}$$

where $h_j^{(l)}$ is the jth hidden unit at the lth layer ($h_i^{(0)} = x_i$ is the ith dimension of input feature), $b_j^{(l)}$ is the bias of the jth
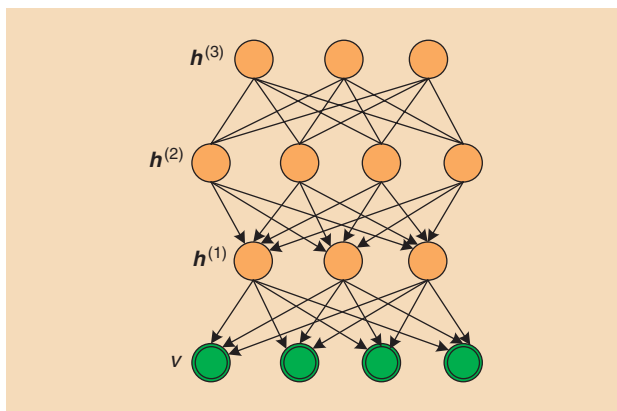
[FIG8] The graphical model representation for a three-hidden-layer DBN.

unit at the lth layer, and $w_{ij}^{(l)}$ is the weight associated with the link from $h_i^{(l-1)}$ to $h_j^{(l)}$. The form of activation functions at the output layer depends on the task. For multiclass classification tasks, a softmax function is typically used

$$\tilde{y}_j = \frac{\exp\left\{b_j^{(L+1)} + \sum_i h_i^{(L)} w_{ij}^{(L+1)}\right\}}{\sum_k \exp\left\{b_k^{(L+1)} + \sum_i h_i^{(L)} w_{ik}^{(L+1)}\right\}}, \tag{40}$$

where $\tilde{y}_j = h_j^{(L+1)}$ gives the posterior probability of the jth class and L is the number of hidden layers. For regression tasks, a linear activation function is often used

$$\tilde{y}_j = b_j^{(L+1)} + \sum_i h_i^{(L)} w_{ij}^{(L+1)}. \tag{41}$$

The set of parameters of an L-hidden-layer DNN consists of $\lambda = \{\boldsymbol{b}^{(1)}, W^{(1)}, \ldots, \boldsymbol{b}^{(L+1)}, W^{(L+1)}\}$. They can be optimized in a supervised way by minimizing a loss function that measures the difference between data and predicted outputs using the backpropagation algorithm [81]. For classification tasks, the cross entropy between correct and predicted class posterior probabilities is often used as the loss function

$$\mathcal{L}(\boldsymbol{y}, \tilde{\boldsymbol{y}}; \lambda) = -\sum_j y_j \log(\tilde{y}_j), \tag{42}$$
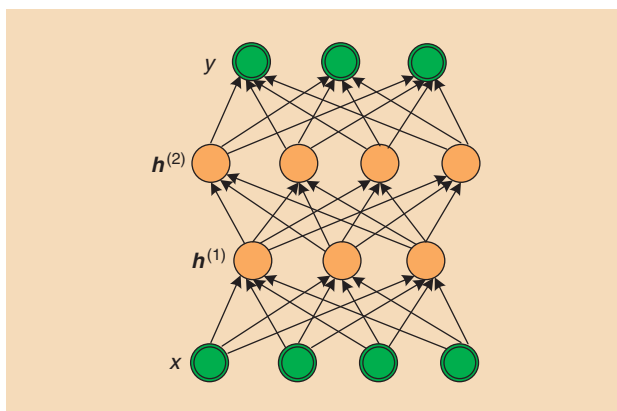
[FIG9] The model representation for a two-hidden-layer DNN.

where y_j denotes the correct class posterior probability given input, which is typically a binary value. For regression tasks, the mean square error is commonly adopted as the loss function

$$\mathcal{L}(y, \tilde{y}; \lambda) = \sum_j (y_j - \tilde{y}_j)^2, \qquad (43)$$

where y_j and $\tilde{y}_j$ are the jth dimension of the correct and predicted outputs, respectively. A DNN for regression can be considered a probabilistic model representing a conditional PDF of y given x using a Gaussian distribution, i.e.,

$$p(y \mid x, \lambda) = \mathcal{N}(y; \tilde{y}, I), \qquad (44)$$

where I is an identity matrix and $\tilde{y}$ depends on x and λ. Thus, minimizing the mean square error between $\tilde{y}$ and y with respect to λ is equivalent to the ML estimation of λ.

DNNs can be powerful models of the highly complex and nonlinear relationship between inputs and outputs. However, it is difficult to train a DNN with many hidden layers. The error signal in back-propagation training decays as it is back-propagated along many hidden layers, which leads to the vanishing gradient problem [82], i.e., the lower layers cannot get much information about how to update their model parameters. Supervised training of DNNs can also result in overfitting to training data because of the power of DNNs to represent training samples. To avoid this problem, unsupervised pretraining techniques, which use DBN (stacked RBMs) weights to initialize a DNN, were proposed [17]. To build an L-hidden-layer DNN, an L-hidden-layer DBN is first trained. Then, weights of the DBN are used to initialize the weights of the DNN. After initializing the DNN weights, supervised fine-tuning is conducted using back-propagation to adjust the weights estimated in pretraining. This unsupervised pretraining strategy can provide a better starting point for supervised fine-tuning than random initialization and reduce overfitting significantly.

Besides RBMs, autoencoders (AEs) are another form of model that can be used for pretraining DNNs in a layerwise manner. An AE is a particular type of one-hidden-layer neural network [83]. It first maps an input vector x to a hidden representation h using a weight matrix W and then maps h back into a reconstruction $\tilde{x}$ of the same shape as x using a weight matrix W'. The two weight matrices may optionally be constrained: $W' = W^\top$. The parameters are optimized such that the average reconstruction error from x to $\tilde{x}$ is minimized. The reconstruction error can be measured using either the mean square error or the cross-entropy criterion depending on the assumed distribution on the input features.

To prevent the hidden layer from simply learning the identity transform, a common modification of the AE is the DAE [21], which is trained to reconstruct the original input from a corrupted copy. Compared with RBMs, one of the advantages of using AEs and DAEs is that many traditional optimization algorithms for neural networks can be used in training. The DAE can also be stacked to form a particular type of DNN, called a deep DAE, through unsupervised pretraining and supervised fine-tuning. While pretraining each layer, the hidden representations

given by the DAE of the layer below are used as the input to the current layer. For supervised fine-tuning, an output layer is added on top of the network and the weights of the entire network are adjusted to minimize the cost function [83].

ACOUSTIC MODELING USING DEEP LEARNING TECHNIQUES FOR SPSG

Given the success of applying deep learning to a variety of speech tasks, we believe that the approach can also be applied to acoustic speech modeling in speech generation to overcome the limitations mentioned earlier and to achieve better input-to-cluster and/or cluster-to-feature mapping. Applications of the deep learning techniques to SPSG had not been investigated until very recently. During the last year, several articles on the topic for speech synthesis [23]–[26], [33], [34], voice conversion [27]–[29], and speech enhancement [30]–[32] have been published. They reported positive results that the deep learning techniques improved the naturalness, similarity, and/or quality of generated speech. These deep learning approaches can be classified into three categories according to the modeling steps, as well as the relationship between the input and output features represented in the model.

CLUSTER-TO-FEATURE MAPPING USING DEEP GENERATIVE MODELS

In this approach, the deep learning techniques are applied to the cluster-to-feature mapping step of acoustic modeling for SPSG, i.e., to describe the distribution of acoustic features at each cluster. The input-to-cluster mapping, which determines the clusters from the input features, still uses conventional approaches, such as decision trees and state-duration PDFs in HMM-based speech synthesis and posterior probabilities of mixture components in GMM-based voice conversion. One example of this approach is HMM-based speech synthesis using RBMs and DBNs for spectral modeling [24]. This work improves the conventional spectral modeling approach in HMM-based parametric speech synthesis. Improvement was achieved in two aspects: First, raw spectral envelopes extracted by speech transformation and representation based on adaptive interpolation of weighted spectrum (STRAIGHT) analysis [84] rather than the low-dimensional representations, such as MCCs or line spectral pairs (LSPs) derived from these spectral envelopes, were modeled. Second, RBMs and DBNs were adopted to replace single Gaussian distributions at the leaf nodes of decision trees. The model structure of this approach is shown in Figure 10. To simplify model training with high-dimensional spectral features, decision trees and state alignments were assumed to be given.

At the acoustic feature extraction stage using STRAIGHT analysis, original spectral envelopes were stored in addition to spectral parameters. The context-dependent HMMs for low-dimensional spectral parameters and F_0 features were estimated according to the approach introduced in the section "HMM-Based Speech Synthesis." A single Gaussian distribution was used to model the spectral parameters at each leaf node of the decision trees. Then, a state-level forced alignment was carried out with the trained HMMs. The state boundaries obtained were used to gather the

spectral envelopes for each decision tree's leaf node. Then, an RBM or a DBN was trained at each leaf node according to the ML criterion. In this approach, the spectral envelope features at each frame consisting of static, velocity, and acceleration components correspond to the visible vector v in (23) for RBMs and (38) for DBNs.

To simplify model estimation, each dimension of the spectral envelope features was normalized to zero mean and unit variance before training RBMs or DBNs, and the variance parameters σ_i^2 in (25) were fixed to 1 for each leaf node. As a result, a set of context-dependent RBM-HMMs or DBN-HMMs is trained for modeling the spectral envelopes.

At synthesis time, the speech parameter-generation algorithm was used to generate the spectral envelopes. The optimal sequences of spectral envelopes were determined so as to maximize their output probability given the RBM-HMM or the DBN-HMM. If a single Gaussian distribution is adopted as the state-output PDFs of HMMs, and the state sequence is given, there is a closed-form solution to determine the optimal acoustic feature trajectories [7]. However, the marginal PDFs of RBMs and DBNs are much more complicated than a single Gaussian distribution. Thus, there is no closed-form solution to find the optimal acoustic feature trajectories. To avoid this problem, a Gaussian approximation was applied before the parameter-generation stage as a simplification. At each decision tree leaf node of decision trees, a Gaussian distribution $\mathcal{N}(v; \boldsymbol{\mu}, \Sigma)$ was constructed, where

$$\boldsymbol{\mu} = \arg \max_v p(v \mid \lambda) \qquad (45)$$

was the mode vector estimated [24] from $p(v \mid \lambda)$ for each RBM or DBN and Σ was a diagonal covariance matrix computed from the training samples associated with the leaf node. Because each dimension of the training samples of v was normalized to zero mean and unit variance, a denormalization processing was conducted before parameter-generation to derive the distributions of the original spectral envelope features from the estimated $\boldsymbol{\mu}$ and Σ. The RBMs/DBNs at the leaf nodes were replaced by these Gaussian distributions at the synthesis stage. Therefore, the speech parameter-generation algorithm can be followed to predict the spectral envelopes. For details about the mode estimation algorithm, refer to [24].

A group of subjective evaluations has been conducted to prove the effectiveness of this approach [24]. Some evaluation results are summarized and shown in Table 1. In this table, each line presents the preference percentages given by a preference listening test conducted between two systems. For example, the first row means that 48% of the stimuli generated by the GMM system was judged by the listeners to be better than those of the baseline system. The percentage of converse preference was 18.67%. The baseline system was constructed using Mel-cepstra and single Gaussian distributions for cluster-to-feature mapping. At training time, Mel-cepstra were derived from the spectral envelopes extracted by STRAIGHT. At synthesis time, the spectral envelopes recovered from the generated mel-cepstra were sent into STRAIGHT to reconstruct speech waveforms. A system using spectral envelopes and single Gaussian distributions for cluster-to-feature mapping
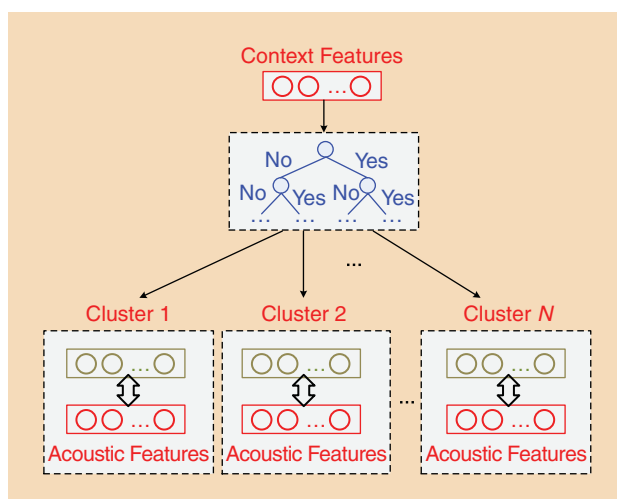
[FIG10] A model structure of cluster-to-feature mapping using RBMs for HMM-based speech synthesis [24].

was also constructed. However, it was found that this system had very similar synthetic results to the baseline system. Some detailed explanation can be found in [24], which means that simply replacing mel-cepstra with spectral envelopes is not helpful if the model structures are not modified accordingly. Therefore, the baseline system was adopted as a representative for these two systems in the subjective evaluation to simplify the test design. The GMM and RBM systems adopted GMMs of eight mixtures and RBMs of 50 hidden units to model the distribution of spectral envelopes at each leaf node of the decision trees. No postfiltering techniques, such as GV-based parameter-generation [13], were applied to any of these systems. It can be seen from the table that the use of RBMs to model the spectral envelopes at each leaf node achieved significantly better naturalness than the use of single Gaussian distributions and GMMs. A comparison between the spectral envelopes generated by the baseline system and the RBM system is shown in Figure 11. From this figure, we can observe the enhanced formant structures after modeling the spectral envelopes using RBMs.

In addition to speech synthesis, this approach was also applied to other speech generation tasks, such as voice conversion [27]. Similar to conventional GMM-based voice conversion, the input-to-cluster mapping in [27] was determined by the posterior probabilities of mixture components of a trained GMM, given the input acoustic features. Then, RBMs were adopted to model the joint PDFs between the source and target acoustic features for each cluster. The subjective evaluation results also demonstrated the effectiveness of

[TABLE 1] THE SUBJECTIVE PREFERENCE SCORES (%) AMONG SPEECH SYNTHESIZED USING THE BASELINE, GMM, AND RBM SYSTEMS.

BASELINE	GMM	RBM	N/P*	p
18.67	**48**	–	33.33	0.0014
5.33	–	**70.67**	24	0
–	16	**69.33**	14.67	0

* N/P denotes "no preference" [24].
The systems that achieved significantly better preference at the p < 0.05 level are in bold font.

this approach when either MCCs or spectral envelopes were used as spectral features. The mean opinion score (MOS) of similarity of the converted speech improved from 2.83 to 3.13, and the MOS of naturalness increased from 2.90 to 3.45, respectively [27].

INPUT-TO-FEATURE MAPPING USING DEEP JOINT MODELS

This approach uses a single deep generative model to achieve the integrated input-to-feature mapping by modeling the joint PDF between the input and output features. For example, a synthesis method using a multidistribution DBN (MD-DBN) has been proposed in [25] with input features capturing linguistic contexts and output features being acoustic features. More specifically, the input contextual features for speech synthesis were the tonal syllables in Mandarin Chinese, which were encoded within a 1-of-k code following the categorical distribution (i.e., the generalized Bernoulli distribution). The output acoustic features to be generated consisted of syllable-level spectrum and excitation features. Each syllable was represented by an acoustic feature supervector, which consisted of multiple frames of Mel-generalized cepstral coefficients (MGCs), log-energy, $\log F_0$, and voiced/unvoiced (U/V) flags. These frames were uniformly spaced within the boundary of a syllable. Different types of acoustic features including spectrum and excitation parameters are modeled by a single network so that the correlation between them can be modeled. Syllable duration was modeled and predicted separately in this framework.

To model the different distributions of the binary data (i.e., the U/V flags) and the continuous data (i.e., the MGCs and $\log F_0$), the approach used an MD-DBN, as shown in Figure 12. This consisted of the building blocks of RBMs, with different types of distribution units in the visible layer. Gaussian distributions were used for the spectral data and $\log F_0$, and Bernoulli distributions for the U/V flags, to form the Gaussian–Bernoulli RBM (GB-RBM) for the bottom layer. Training of the MD-DBN began with unsupervised

learning, where an MD-DBN with $L-1$ hidden layers was first trained using the acoustic features as observations as shown in the right part of Figure 12. The MD-DBN was built by stacking up multiple Bernoulli RBMs (B-RBMs) on top of the bottom GB-RBM layer; thus, the depth of the model could be easily controlled. This was followed by supervised learning where the $(L-1)$th layer was extended with a 1-of-k vector x that encoded Mandarin syllable IDs and then learned one more layer on top. This additional layer modeled the joint distribution between the syllable IDs and the hidden activations of the supervector using the Categorical-Bernoulli RBM (CB-RBM).

This training paradigm has three advantages over HMM-based synthesis: 1) It models all training data in a centralized network and avoids data partitioning. Instead of using thousands of Gaussian distributions to piece the acoustic space together as in the HMM-based approach, this approach uses only one MD-DBN to portray the whole acoustic space, which potentially reduces the requirements of training data and increases the efficiency of model parameters. 2) The supervector consists of multiple acoustic frames from a syllable with temporal dynamics intact, which can be captured by the MD-DBN. This differs from the HMM-based synthesis, which assumes that acoustic observations are dependent only on the current hidden state. Since the correlations in the temporal domain can be captured directly by the MD-DBN, the use of dynamic features can be eliminated. 3) In the frequency domain, the correlations between spectral coefficients within a single frame can also be modeled by the MD-DBN, which does not adopt any independence assumptions such as those introduced by the use of a GMM with a diagonal covariance matrix. As a result, the decoupling process in the speech feature extraction can be eliminated to preserve more information.

At synthesis time, the contextual features x were first determined for each syllable by text analysis. Then, alternative Gibbs sampling using $P(h_i^{(L)}=1 \mid x, h^{(L-1)}, \lambda)$ and $P(h_j^{(L-1)}=1 \mid h^{(L)}, \lambda)$ were conducted with the x clamped to update $h^{(L-1)}$ until convergence or a maximum number of iterations was reached. Then, the acoustic feature supervector was predicted as the mean vector of $p(y \mid h^{(1)}, \lambda)$, which was determined by recursively generating hidden variables from $h^{(L-1)}$ to $h^{(1)}$. Finally, the generated acoustic features were interpolated according to the predicted syllable durations and were sent into the Mel log spectrum approximation filter [85] to reconstruct the speech waveforms. No postfiltering or global-variance-based voice enhancement techniques were incorporated.

It is worth noting that this acoustic modeling method discarded HMMs and modeled the joint PDF between the input contextual features and the output acoustic features using one single MD-DBN without the conventional two-step mapping. Table 2 shows the five-point Likert scale MOSs of the HMM baseline (HMM), the system predicting MGCs using the proposed MD-DBN approach [DBN (MGCs)], and the system predicting both MGCs and $\log F_0$ using the MD-DBN approach [DBN (MGCs + $\log F_0$)] [25]. Comparing DBN (MGCs) with HMM, we can see that the proposed MD-DBN approach outperforms the conventional HMM baseline for modeling and
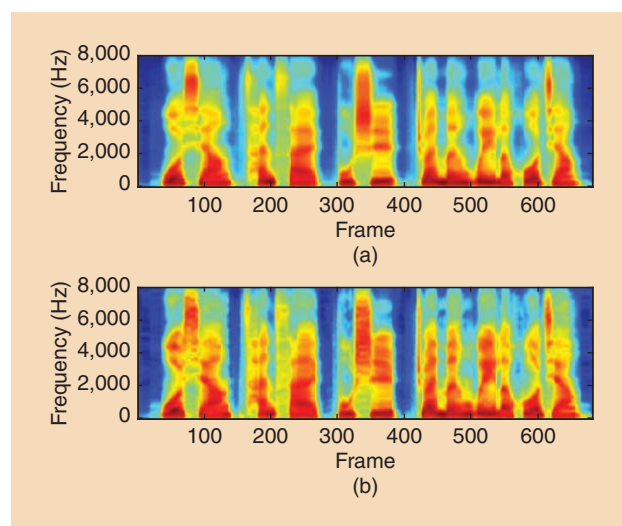
[FIG11] The spectrograms of a segment of synthetic speech using (a) the baseline system and (b) the RBM system [24].

predicting spectral features. The quality degradation from DBN(MGCs) to DBN(MGCs + $\log F_0$) suggests that the low-dimensional F_0 features are not well modeled when combined with high-dimensional spectrum features.
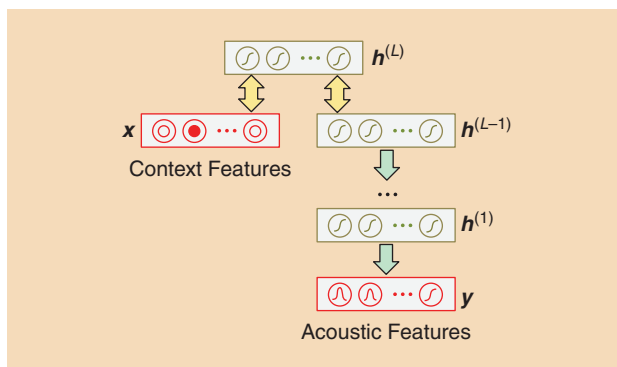
INPUT-TO-FEATURE MAPPING USING DEEP CONDITIONAL MODELS

Similar to the previous approach, this one predicts acoustic features from inputs using an integrated deep generative model. The difference is that this approach models a conditional PDF of output acoustic features, given input features instead of their joint PDF.

A DNN-based speech synthesis approach was proposed in [26]. In this approach, context and acoustic features were treated as inputs and targets of a DNN, respectively, as shown in Figure 13. As introduced in the DNNs section, a DNN describes a conditional PDF of outputs given inputs using a Gaussian distribution. A text to be synthesized was first converted to a sequence of frame-level linguistic context features. The linguistic context features at each frame included binary answers to questions about contexts, numeric context descriptors, position of the current frame within a segment, and segment durations. The acoustic features at each frame were composed of MCCs, $\log F_0$, excitation aperiodicities, their derived dynamic components [3], and binary U/V decisions. The weights of the DNN were trained from pairs of inputs and targets extracted from training data. Like the DBN-based approach discussed in the section "Input-to-Feature Mapping Using Deep Joint Models," as acoustic features include both spectral and excitation parameters and a single DNN is trained, correlations between them can be modeled. At synthesis time, phoneme durations were first determined by a duration prediction module; then, frame-level linguistic context features were composed. By feeding the composed linguistic context features to the trained DNN, output acoustic features were predicted. By using these predicted output acoustic features as means along with the frame-independent variances of output acoustic features computed from all training data, the speech parameter-generation algorithm [7] generated the smooth acoustic feature trajectories. The generated acoustic feature parameters were post-processed by a postfilter (in the experiment reported in [26], postfiltering in the mel-cepstral domain [86] was applied to emphasize formant structure) and then sent to a vocoder to reconstruct a speech waveform.

A subjective preference listening test was conducted to compare the performance of the DNN-based systems with HMM-based systems [26]. The experimental results are shown in Table 3. In this experiment, HMM- and DNN-based systems with similar numbers of parameters were compared. The α in the first column of Table 3 is the scaling factor for the penalty term in the minimum description length (MDL) criterion, which is often used to control the number of parameters in HMM-based systems. It can be seen in the table that, for all three model sizes, the DNN-based system achieved better naturalness than the HMM-based system according to the p values given by hypothesis tests.

Other approaches of DNN-based TTS can be found [33], [34]. These include a hybrid approach between DNN and Gaussian
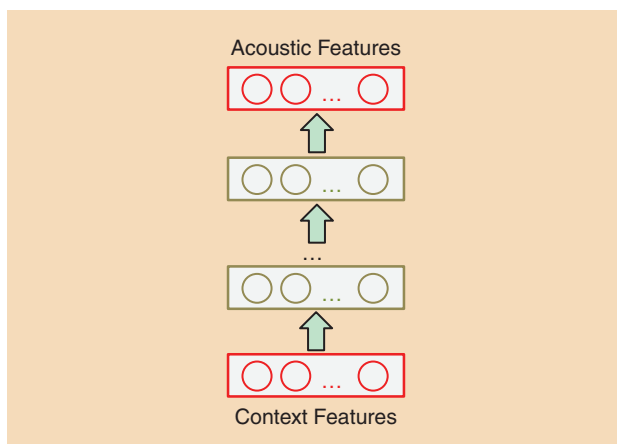
[FIG12] The model structure of input-to-feature mapping using DBN for speech synthesis [25].

[TABLE 2] THE SUBJECTIVE EVALUATION RESULTS FOR THE DBN-BASED SPEECH SYNTHESIS [25].

SYSTEM	MOS
HMM	2.86
DBN (MGCs)	3.09
DBN (MGCs + $\log F_0$)	2.88

process (GP)-based regression [33] to predict $\log F_0$; a DNN that maps linguistic context features to $\log F_0$ was first trained, and then the activations at the last hidden layer were used as inputs for GP-based nonparametric regression. This approach combined the parametric and nonparametric regression models. An alternative approach [34] used a vector-space representation of input texts as inputs of DNN-based TTS. This vector-space representation was derived without using any linguistic resources; only orthographic information (graphemes) was used; thus, it did not require any language knowledge to build a model.

The acoustic modeling approach using deep conditional models has also been applied to other speech generation tasks, such as voice conversion [28], [29] and speech enhancement [30]–[32]. A DNN-based voice conversion approach has been proposed in [28]. In this approach, acoustic features of a source

[FIG13] A model structure of input-to-feature mapping using DNN for speech synthesis [26].

[TABLE 3] THE SUBJECTIVE PREFERENCE SCORES (%) BETWEEN SPEECH SAMPLES FROM THE HMM- AND DNN-BASED SYSTEMS [26].

HMM (α)	DNN (# LAYERS × # UNITS)	N/P	p
15.8 (16)	**38.5** (4 × 256)	45.7	$< 10^{-6}$
16.1 (4)	**27.2** (4 × 512)	56.8	$< 10^{-6}$
12.7 (1)	**36.6** (4 × 1,024)	50.7	$< 10^{-6}$

The systems that achieved significantly better preference at the $p < 0.01$ level are shown in bold font.

voice were mapped to those of a target voice using a DNN that was initialized by concatenating two DBNs. CRBMs have also been used to construct conditional models for voice conversion. In [29], a CRBM was estimated to model a conditional PDF of acoustic features of a target voice given acoustic features from a source voice. For speech enhancement, conditional generative model-based approaches have been proposed for mapping acoustic features extracted from noisy speech to those of clean speech using DNNs [32] or DAEs [30], [31].

COMPARISONS AMONG THESE THREE APPROACHES

The cluster-to-feature mapping approach using deep generative models has the model structure most similar to conventional HMM- or GMM-based approaches. The input-to-cluster mapping step is preserved, and few modifications to the existing speech generation engines are necessary after off-line model training [24]. The input-to-feature mapping approaches using deep joint models or deep conditional models integrate the two-step mapping of acoustic modeling into a single step [25], [26], which can express complicated mapping functions more efficiently and provide better generalization than the approaches using conventional input-to-cluster mapping, such as

decision trees and GMM posterior probabilities. Compared with the sampling-based parameter-generation from a DBN [25], generating acoustic features from a DNN is more straightforward [26]. However, the conditional PDF represented by a DNN is relatively simple because it is a Gaussian distribution with an identity covariance matrix as described in the "DNNs" section. Table 4 summarizes the recently proposed acoustic modeling approaches using deep learning techniques for SPSG. Some discussions on these approaches will be given in the "Discussion" section.

DISCUSSION

PERFORMANCE OF RBMs AS DENSITY MODELS

RBMs are the basis of many deep models such as DBNs and DNNs. As introduced in the "RBMs" section, RBMs have some good properties in describing the distribution of high-dimensional observations with cross-dimension correlations. The performance of GMMs and RBMs in modeling the distribution of mel-cepstra and spectral envelopes for a specific context-dependent HMM state was investigated in [23]. Spectral envelopes were extracted by STRAIGHT analysis [84], and MCCs were derived from the spectral envelopes at each frame. In the experiment, a leaf node with 720 frames was used; 520 frames were used for training and the remaining 200 frames were used as a test set. The number of mixture components in a GMM varied from 1 to 32, and the number of hidden units in an RBM varied from 1 to 1,000. The average log probabilities on the training and test sets for different model structures are shown in Table 5 for MCCs and the spectral envelopes, respectively. It can be seen from the tables that the GMMs overfit more to the training data as the model complexity increased. On the other hand, the RBMs consistently gave good generalization ability even with a large number of hidden units. It can be seen

[TABLE 4] A SUMMARY OF THE PROPOSED ACOUSTIC MODELING APPROACHES USING DEEP LEARNING TECHNIQUES FOR SPSG.

	TASK	MODEL STRUCTURE	INPUT FEATURES	GENERATED ACOUSTIC FEATURES
LING ET AL. 2013 [24] *	SPEECH SYNTHESIS	RBM/DBN-HMM	RICH CONTEXT FEATURES	SPECTRAL ENVELOPES
KANG ET AL. 2013 [25] @	SPEECH SYNTHESIS	DBN	SIMPLE LINGUISTIC FEATURES	MCCs, $\log F_0$, AND U/V
ZEN ET AL. 2013 [26] §	SPEECH SYNTHESIS	DNN	RICH LINGUISTIC CONTEXT FEATURES	MCCs, $\log F_0$, APERIODICITIES, AND U/V
LU ET AL. 2013 [34] §	SPEECH SYNTHESIS	DNN	VECTOR SPACE REPRESENTATION OF TEXTS	LSPs, $\log F_0$, AND APERIODICITIES
FERNANDEZ ET AL. 2013 [33] §	SPEECH SYNTHESIS	DNN-GP	RICH LINGUISTIC CONTEXT FEATURES	$\log F_0$
CHEN ET AL. 2013 [27] *	VOICE CONVERSION	MIXTURE OF RBMs	SPECTRAL ENVELOPES OF SOURCE VOICE	SPECTRAL ENVELOPES
NAKASHIKA ET AL. 2013 [28] §	VOICE CONVERSION	DNN	MCCs OF SOURCE VOICE	MCCs OF TARGET VOICE
WU ET AL. 2013 [29] §	VOICE CONVERSION	CRBM	MCCs OF SOURCE VOICE	MCCs OF TARGET VOICE
LU ET AL. 2013 [30] §	SPEECH ENHANCEMENT	DEEP DAE	POWER SPECTRA OF NOISY SPEECH	POWER SPECTRA OF CLEAN SPEECH
XIA ET AL. 2013 [31] §	SPEECH ENHANCEMENT	DAE	POWER SPECTRA OF NOISY SPEECH	POWER SPECTRA OF CLEAN SPEECH
XU ET AL. 2014 [32] §	SPEECH ENHANCEMENT	DNN	POWER SPECTRA OF NOISY SPEECH	POWER SPECTRA OF CLEAN SPEECH

*, @, and § denote the three categories descibed in the section "Acoustic Modeling Using Deep Learning Techniques for SPSG."
* denotes cluster-to-feature mapping using deep generative models.
@ denotes input-to-feature mapping using deep joint models.
§ denotes input-to-feature mapping using deep conditional models.

from Table 5 that the best GMM and the best RBM had very close test-set log probabilities while modeling the MCCs. However, the RBMs gave much higher test-set log probabilities than the GMMs as shown in Table 5(b). These results can be attributed to the fact that mel-cepstral analysis decorrelates spectral parameters, whereas the advantage of RBMs is to analyze the latent patterns embedded in the high-dimensional raw data with strong interdimensional correlations, such as raw spectral envelopes.

INPUT AND TARGET FEATURES

In acoustic modeling for SPSG, the forms of input features are task dependent. The same is true for the acoustic modeling using deep learning techniques. As shown in Table 4, simple to rich linguistic context features are typically used as input features for speech synthesis [24]–[26], [33], whereas vector-space representation of input texts has also been used [34]. Input features for voice conversion are typically spectral features extracted from a source voice [27]–[29]. Likewise, input features for speech enhancement are typically power spectra extracted from noisy speech [30]–[32].

Various output acoustic features for speech generation have been used, as listed in Table 4. The discussion in the section "Performance of RBMs as Density Models" shows that RBMs and other deep generative models are good at modeling the distribution of high-dimensional acoustic features with cross-dimensional correlations. Thus, some approaches took this into account when selecting their output acoustic features. The cross-dimensional correlations represented by the deep generative models exist in both the frequency domain, e.g., by using raw power spectra or spectral envelopes at each frame [24], [30]–[32], and the temporal domain, e.g., by concatenating the acoustic features of multiple frames [25]. In some speech generation tasks, such as speech synthesis, F_0 is another important acoustic feature to be predicted in addition to spectral parameters. F_0 together with other excitation-related acoustic features, including U/V decisions and aperiodicity ratios, has also been used as a part of target features in some deep-learning-based acoustic modeling approaches [25], [26]. However, the prediction performance of $\log F_0$ was not as good as that of spectral features as shown in the experimental results in [25] and [26].

MODEL STRUCTURES AND MODEL TRAINING

As shown in Table 4, different model structures have been adopted in these approaches. RBMs and DBNs were used to represent joint PDFs and to achieve cluster-to-feature [24], [27] or input-to-feature mapping [25]. On the other hand, DNNs, CRBMs, and DAEs were adopted to represent conditional PDFs and to achieve direct input-to-feature mapping [26], [30]–[32]. The depth of architecture, i.e., the number of hidden layers, is an important characteristic of a deep model. In DBN–HMM-based speech synthesis [24], the experimental results in Table 1 show that increasing the number of layers did not improve the naturalness of synthetic speech because of the difficulty of estimating the mode of a DBN.

In other works [25], [26], [30], [32], the number of hidden layers was tuned to minimize the mean squared error between targets (data) and outputs (predicted acoustic features) on development sets. The results show that multiple hidden layers could achieve better prediction accuracy than a single hidden layer. However, the optimal depth is commonly not as deep as that used in DNN-HMM-based ASR. It is reasonable considering that the amount of training data for speech generation tasks is limited compared with ASR. In the DNN-based approaches, different initialization strategies have been employed, e.g., random initialization for speech synthesis [26], structured pretraining using DBNs and NNs for voice conversion [28], and pretraining using stacked AEs or RBMs for speech enhancement [30], [32]. Considering the heavy computational cost of training RBMs, DNNs, and other deep models, graphics processing unit-based acceleration was applied to reduce the training time [25], [26].

A COMPARISON BETWEEN SPEECH SYNTHESIS AND RECOGNITION BOTH USING DNN-HMMS

The DNN-HMM is the dominant form of acoustic modeling with deep structures for ASR [22]. In this approach, a DNN is trained to map input acoustic features (e.g., mel-frequency cepstral coefficients, log-filterbank features, etc.) to posterior probabilities of leaf nodes of decision trees at each frame. HMMs are used to connect the hidden states with the higher-level linguistic representations for decoding with language models at recognition time. While there seems to be a converging deep learning

[TABLE 5] THE AVERAGE LOG PROBABILITIES ON THE TRAINING AND TEST SETS WHEN MODELING THE MEL-CEPSTRA AND SPECTRAL ENVELOPES OF A SPECIFIC STATE USING DIFFERENT MODELS [23].

MEL-CEPSTRA COEFFICIENTS

| | AVERAGE LOG PROBABILITY | | NUMBER OF PARAMETERS |
	TRAIN	TEST	
GMM (1)-DIAG	−58.176	−56.380	82
GMM (4)-DIAG	−51.188	−53.097	328
GMM (16)-DIAG	−40.869	−59.492	1,312
GMM (32)-DIAG	−29.973	−72.056	2,624
GMM (1)-FULL	−30.883	−54.648	902
RBM (1)	−56.464	−55.244	83
RBM (10)	−52.416	−52.660	461
RBM (50)	−51.840	−53.636	2,141
RBM (200)	−53.554	−55.020	8,441
RBM (1,000)	−55.797	−56.940	42,041

SPECTRAL ENVELOPES

| | AVERAGE LOG PROBABILITY | | NUMBER OF PARAMETERS |
	TRAIN	TEST	
GMM (1)-DIAG	−727.915	−728.647	1,026
GMM (4)-DIAG	−599.642	−648.818	4,104
GMM (16)-DIAG	−485.072	−665.609	16,416
GMM (32)-DIAG	−379.980	−717.523	32,832
GMM (1)-FULL	2,207.177	−89,202.438	132,354
RBM (1)	−685.799	−700.938	1,027
RBM (10)	−629.906	−649.823	5,653
RBM (50)	−587.146	−628.222	30,317
RBM (200)	−576.461	−617.480	103,313
RBM (1,000)	−562.439	−583.169	514,513

The numbers in the brackets indicate the Gaussian mixture numbers for the GMMs and the hidden unit numbers or the RBMS. "DIAG" and "FULL" denote using diagonal and full covariance matrices, respectively.

architecture based on the DNN-HMM for the dominant use in ASR, there has been a greater variety of model structures proposed for SPSG using deep learning techniques, where the variety can be seen in Table 4. Among them, the DNN-based conditional modeling approach [26], [32] adopts a model structure quite similar to the DNN-HMM for acoustic modeling in ASR. One main difference is in the activation functions used at the DNN's output layers: the softmax layer for multiclass classification in ASR versus the linear layer for regression in SPSG.

In DNN–HMM-based ASR, acoustic features are the input to a DNN for classification, while DNN-based SPSG predicts acoustic features for speech generation. Therefore, the acoustic features used in DNN-based SPSG should take into account the requirement of reconstructing speech waveforms. Some acoustic features that are not adopted in DNN–HMM-based ASR, such as excitation-related features [26] and power spectra [32], have been used in DNN-based SPSG.

CONCLUSIONS

This article provides an overview of the emerging speech generation approaches using deep learning techniques. Compared with the conventional acoustic modeling methods in SPSG based on the use of HMMs and GMMs, deep joint models (e.g., RBMs and DBNs) and deep conditional models (e.g., CRBMs and DNNs), which we reviewed in this article, are better able to describe the complex and nonlinear relationship between the inputs and targets of the SPSG system and, therefore, improve the naturalness, similarity to the target speaker, and quality of the generated speech. Various implementations of building acoustic models using deep learning for SPSG in the current literature have been reviewed and compared. To facilitate a review of the area and to offer insights into the different approaches reported in the literature, we categorize them into three classes, describe and analyze each, and make connections in a systematic manner.

Despite the empirical successes of a range of deep learning methods in SPSG as reviewed in this article, there remain important issues that need further investigation to make full use of the intrinsic strength of deep learning models and methods in SPSG. For example, current attempts have not achieved positive results in modeling and prediction of F_0 using deep generative models [25], [26]. Considering the different physiological mechanisms between the production of F_0 and of spectral features, deep model structures designed specifically for F_0 modeling and prediction may be necessary. Furthermore, few considerations have been made thus far in deep learning approaches to model the temporal dependencies among the sequence of acoustic features. We believe that a promising direction to pursue in the near future is to apply the deep generative models with better temporal modeling abilities, such as recurrent neural networks, to the SPSG tasks in the future.

AUTHORS

Zhen-Hua Ling (zhling@ustc.edu.cn) received his B.E. degree in electronic information engineering and his M.S. and Ph.D. degrees in signal and information processing from the University of Science and Technology of China, Hefei, in 2002, 2005, and 2008, respectively. From 2007 to 2008, he was a Marie Curie Fellow at the Centre for Speech Technology Research, University of Edinburgh, United Kingdom. From 2008 to 2011, he was a joint postdoctoral researcher at the University of Science and Technology of China and iFLYTEK Co., Ltd., China. He is currently an associate professor at the University of Science and Technology of China. He was also with the University of Washington, United States, as a visiting scholar from 2012 to 2013. His research interests include speech processing, speech synthesis, voice conversion, speech analysis, and speech coding. He received the 2010 IEEE Signal Processing Society Young Author Best Paper Award. He is a Member of the IEEE.

Shi-Yin Kang (sykang@se.cuhk.edu.hk) received his B.S. degree in automation and his M.S. degree in computer science and technology from Tsinghua University, Beijing, China, in 2007 and 2010, respectively. He is currently pursuing his Ph.D. degree in the Department of Systems Engineering and Engineering Management of The Chinese University of Hong Kong, with a research focus in speech synthesis. His current research interests include statistical parametric speech synthesis and applications in machine learning.

Heiga Zen (heigazen@google.com) received his Ph.D. degree from the Nagoya Institute of Technology, Japan, in 2006. Before joining Google in 2011, he was an intern/cooperative researcher at the IBM T.J. Watson Research Center, Yorktown Heights, New York, from 2004 to 2005, and a research engineer at Toshiba Research Europe Ltd. Cambridge Research Laboratory, Cambridge, United Kingdom, from 2008 to 2011. His research interests include speech synthesis and recognition. He was one of the original authors and the first maintainer of the hidden Markov model-based speech synthesis system.

Andrew Senior (andrewsenior@google.com) received his Ph.D. degree from the University of Cambridge, United Kingdom, for his work on handwriting recognition with recurrent neural networks. He worked at the IBM T.J. Watson Research Center, Yorktown Heights, New York, in the areas of fingerprint, face, and audio-visual speech recognition, as well as visual tracking and privacy protection. He is currently a research scientist at Google, where he works on deep learning for speech recognition and synthesis. He coauthored *A Guide to Biometrics* and recently coorganized an International Conference on Machine Learning Workshop on deep learning.

Mike Schuster (schuster@google.com) received his Diplom Ingenieur degree in electrical engineering from Gerhard-Mercator University in Duisburg, Germany, and his Ph.D. degree from the Nara Institute of Science and Technology, Japan. He was with Advanced Telecommunications Research Laboratories in Kyoto, Japan; Nuance in the United States; and Nippon Telegraph and Telephone in Japan, where he mostly worked on various machine-learning techniques applied to sequential modeling and, in particular, speech recognition. He is currently a research scientist at Google working on machine-learning techniques for speech recognition, speech synthesis, translation, recommendation systems, and related areas.

Xiao-Jun Qian (xjqian@se.cuhk.edu.hk) received his B.E. degree in electrical engineering from Fudan University, Shanghai,

China, in 2007. From 2007 to 2010, was with the speech group of Microsoft Research Asia. He joined the Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, as a Ph.D. candidate in 2009. His research interests include discriminative training, subspace acoustic modeling, and deep learning. He was the recipient of the 2010 Microsoft Research Asia Fellowship.

Helen Meng (hmmeng@se.cuhk.edu.hk) received her S.B., S.M., and Ph.D. degrees, all in electrical engineering, from the Massachusetts Institute of Technology. She joined The Chinese University of Hong Kong in 1998, where she is currently a professor and chair in the Department of Systems Engineering and Engineering Management. She was also the associate dean of research of the Faculty of Engineering between 2005 and 2010. Her research interest is in the area of human–computer interaction via multimodal and multilingual spoken language systems, speech retrieval technologies, and computer-aided pronunciation training. She served as the editor-in-chief of *IEEE Transactions on Audio, Speech, and Language Processing* from 2009 to 2011. She is an elected board member of the International Speech Communication Association as well as an elected member of the IEEE Signal Processing Society Board of Governors. She is a Fellow of the IEEE.

Li Deng (deng@microsoft.com) received his Ph.D. degree from the University of Wisconsin–Madison. He was a tenured professor from 1989 to 1999 at the University of Waterloo, Ontario, Canada, and then joined Microsoft Research, Redmond, Washington, where he is currently a principal research manager of its Deep Learning Technology Center. Since 2000, he has also been an affiliate full professor at the University of Washington, Seattle, teaching computer speech processing. He has been granted more than 60 U.S. or international patents, and has received numerous awards and honors bestowed by the IEEE, the International Speech Communication Association (ISCA), the Acoustical Society of America (ASA), and Microsoft, including the 2013 IEEE Signal Processing Society Best Paper Award on deep neural networks for speech recognition. He has authored or coauthored four books. He is a Fellow of the ASA, IEEE, and ISCA. He was the editor-in-chief of *IEEE Signal Processing Magazine* from 2009 to 2011 and the editor-in-chief of *IEEE Transactions on Audio, Speech, and Language Processing* from 2012 to 2014. His recent research interests and activities have been focused on deep learning and machine intelligence applied to large-scale text analysis and to speech/language/image multimodal processing.

REFERENCES

[1] K. Tokuda, Y. Nankaku, T. Toda, H. Zen, H. Yamagishi, and K. Oura, "Speech synthesis based on hidden Markov models," *Proc. IEEE*, vol. 101, no. 5, pp. 1234–1252, 2013.

[2] T. Toda, A. Black, and K. Tokuda, "Voice conversion based on maximum-likelihood estimation of spectral parameter trajectory," *IEEE Trans. Audio Speech Lang. Process.*, vol. 15, no. 8, pp. 2222–2235, 2007.

[3] H. Zen, T. Toda, M. Nakamura, and K. Tokuda, "Details of Nitech HMM-based speech synthesis system for the Blizzard Challenge 2005," *IEICE Trans. Inf. Syst.*, vol. E90-D, no. 1, pp. 325–333, 2007.

[4] Z.-H. Ling, Y.-J. Wu, Y.-P. Wang, L. Qin, and R.-H. Wang, "USTC system for Blizzard Challenge 2006: An improved HMM-based speech synthesis method," in *Proc. Blizzard Challenge Workshop*, 2006.

[5] H. Zen, K. Tokuda, and A. Black, "Statistical parametric speech synthesis," *Speech Commun.*, vol. 51, no. 11, pp. 1039–1064, 2009.

[6] T. Yoshimura, K. Tokuda, T. Masuko, T. Kobayashi, and T. Kitamura, "Simultaneous modeling of spectrum, pitch and duration in HMM-based speech synthesis," in *Proc. Eurospeech*, 1999, pp. 2347–2350.

[7] K. Tokuda, T. Yoshimura, T. Masuko, T. Kobayashi, and T. Kitamura, "Speech parameter-generation algorithms for HMM-based speech synthesis," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2000, vol. 3, pp. 1315–1318.

[8] H. Zen, K. Tokuda, and T. Kitamura, "Reformulating the HMM as a trajectory model by imposing explicit relationships between static and dynamic feature vector sequences," *Comput. Speech Lang.*, vol. 21, no. 1, pp. 153–173, 2006.

[9] H. Zen, M. Gales, Y. Nankaku, and K. Tokuda, "Product of experts for statistical parametric speech synthesis," *IEEE Trans. Audio Speech Lang. Processing*, vol. 20, no. 3, pp. 794–805, 2012.

[10] T. Koriyama, T. Nose, and T. Kobayashi, "Statistical parametric speech synthesis based on Gaussian process regression," *IEEE J. Select. Topics Signal Processing*, vol. 8, no. 2, pp. 173–183, 2014.

[11] Y.-J. Wu and R.-H. Wang, "Minimum generation error training for HMM-based speech synthesis," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2006, pp. 89–92.

[12] Y.-J. Wu and K. Tokuda, "Minimum generation error training with direct log spectral distortion on LSPs for HMM-based speech synthesis," in *Proc. Interspeech*, 2008, pp. 577–580.

[13] T. Toda and K. Tokuda, "A speech parameter-generation algorithm considering global variance for HMM-based speech synthesis," *IEICE Trans. Inf. Syst.*, vol. E90-D, no. 5, pp. 816–824, 2007.

[14] T. Tiomkin, D. Malah, and S. Shechtman, "Statistical text-to-speech synthesis based on segment-wise representation with a norm constraint," *IEEE Trans. Audio Speech Lang. Processing*, vol. 18, no. 5, pp. 1077–1082, 2010.

[15] Z.-H. Ling and L.-R. Dai, "Minimum Kullback-Leibler divergence parameter-generation for HMM-based speech synthesis," *IEEE Trans. Audio Speech Lang. Processing*, vol. 20, no. 5, pp. 1492–1502, 2012.

[16] G. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computat.*, vol. 18, no. 7, pp. 1527–1554, 2006.

[17] G. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[18] R. Salakhutdinov, "Learning deep generative models," Ph.D. thesis, Univ. of Toronto, 2009.

[19] P. Smolensky, "Information processing in dynamical systems: Foundations of harmony theory," in *Parallel Distributed Processing*, D. E. Rumelhart and J. L. McClelland, Eds. Cambridge, MA: MIT Press, 1986, vol. 1, ch. 6, pp. 194–281.

[20] L. Deng, M. Seltzer, D. Yu, A. Acero, A. Mohamed, and G. Hinton, "Binary coding of speech spectrograms using a deep auto-encoder," in *Proc. Interspeech*, 2010, pp. 1692–1695.

[21] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, pp. 3371–3408, Dec. 2010.

[22] G. Hinton, L. Deng, D. Yu, G. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Processing Mag.*, vol. 29, no. 6, pp. 82–97, 2012.

[23] Z.-H. Ling, L. Deng, and D. Yu, "Modeling spectral envelopes using restricted Boltzmann machines for statistical parametric speech synthesis," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 7825–7829.

[24] Z.-H. Ling, L. Deng, and D. Yu, "Modeling spectral envelopes using restricted Boltzmann machines and deep belief networks for statistical parametric speech synthesis," *IEEE Trans. Audio Speech Lang. Processing*, vol. 21, no. 10, pp. 2129–2139, 2013.

[25] S.-Y. Kang, X.-J. Qian, and H. Meng, "Multi-distribution deep belief network for speech synthesis," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 8012–8016.

[26] H. Zen, A. Senior, and M. Schuster, "Statistical parametric speech synthesis using deep neural networks," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 7962–7966.

[27] L.-H. Chen, Z.-H. Ling, Y. Song, and L.-R. Dai, "Joint spectral distribution modeling using restricted Boltzmann machines for voice conversion," in *Proc. Interspeech*, 2013, pp. 3052–3056.

[28] T. Nakashika, R. Takashima, T. Takiguchi, and Y. Ariki, "Voice conversion in high-order eigen space using deep belief nets," in *Proc. Interspeech*, 2013, pp. 369–372.

[29] Z.-Z Wu, E.S. Chng, and H.-Z. Li, "Conditional restricted Boltzmann machine for voice conversion," in *Proc. ChinaSIP*, 2013, pp. 104–108.

[30] X. Lu, Y. Tsao, S. Matsuda, and C. Hori, "Speech enhancement based on deep denoising autoencoder," in *Proc. Interspeech*, 2013, pp. 436–440.

[31] B.-Y. Xia and C.-C. Bao, "Speech enhancement with weighted denoising autoencoder," in *Proc. Interspeech*, 2013, pp. 3444–3448.

[32] Y. Xu, J. Du, L.-R. Dai, and C.-H. Lee, "An experimental study on speech enhancement based on deep neural networks," *IEEE Signal Processing Lett.*, vol. 21, no. 1, pp. 65–68, 2014.

[33] R. Fernandez, A. Rendel, B. Ramabhadran, and R. Hoory, "f_0 contour prediction with a deep belief network-Gaussian process hybrid model," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 6885–6889.

[34] H. Lu, S. King, and O. Watts, "Combining a vector space representation of linguistic context with a deep neural network for text-to-speech synthesis," in *Proc. ISCA SSW8*, 2013, pp. 261–265.

[35] S.-Y. Kang and H. Meng, "Statistical parametric speech synthesis using weighted multi-distribution deep belief network," in *Proc. Interspeech,* 2014, pp. 1959–1963.

[36] Y. Qian, Y.-C. Fan, W.-P. Hu, and F. K. Soong, "On the training aspects of deep neural networks (DNN) for parametric TTS synthesis," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 3857–3861.

[37] H. Zen and A. Senior, "Deep mixture density networks for acoustic modeling in statistical parametric speech synthesis," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 3872–3876.

[38] T. Nakashika, T. Takiguchi, and Y. Ariki, "Voice conversion in time-invariant speaker-independent space," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 7939–7943.

[39] L.-H. Chen, Z.-H. Ling, L.-J. Liu, and L.-R. Dai, "Voice conversion using deep neural networks with layer-wise generative training," in *Proc. IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 12, 2014, pp. 1859–1872.

[40] Y. Xu, J. Du, L.-R. Dai, and C.-H. Lee, "Dynamic noise aware training for speech enhancement based on deep neural networks," in *Proc. Interspeech* (to be published).

[41] K. Tokuda, T. Kobayashi, and S. Imai, "Speech parameter-generation from HMM using dynamic features," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 1995, pp. 660–663.

[42] J. Yamagishi, T. Nose, H. Zen, Z.-H. Ling, T. Toda, K. Tokuda, S. King, and S. Renals, "Robust speaker-adaptive HMM-based text-to-speech synthesis," *IEEE Trans. Audio Speech Lang. Processing*, vol. 17, no. 6, pp. 1208–1230, 2009.

[43] J. Yamagishi and T. Kobayashi, "Average-voice-based speech synthesis using HSMM-based speaker adaptation and adaptive training," *IEICE Trans. Inf. Syst.*, vol. E90-D, no. 2, pp. 533–543, 2007.

[44] J. Yamagishi, K. Onishi, T. Masuko, and T. Kobayashi, "Acoustic modeling of speaking styles and emotional expressions in HMM-based speech synthesis," *IEICE Trans. Inf. Syst.*, vol. E88-D, no. 3, pp. 503–509, 2005.

[45] M. Tachibana, J. Yamagishi, T. Masuko, and T. Kobayashi, "Speech synthesis with various emotional expressions and speaking styles by style interpolation and morphing," *IEICE Trans. Inf. Syst.*, vol. E88-D, no. 11, pp. 2484–2491, 2005.

[46] T. Nose, J. Yamagishi, and T. Kobayashi, "A style control technique for HMM-based expressive speech synthesis," *IEICE Trans. Inf. Syst.*, vol. E90-D, no. 9, pp. 1406–1413, 2007.

[47] L. Saheer, J. Dines, and P. N. Garner, "Vocal tract length normalization for statistical parametric speech synthesis," *IEEE Trans. Audio Speech Lang. Processing*, vol. 20, no. 7, pp. 2134–2148, 2012.

[48] H. Zen, N. Braunschweiler, S. Buchholz, M.J.F. Gales, K. Knill, S. Krstulovic, and J. Latorre, "Statistical parametric speech synthesis based on speaker and language factorization," *IEEE Trans. Audio Speech Lang. Processing*, vol. 20, no. 6, pp. 1713–1724, 2012.

[49] Z.-H. Ling, K. Richmond, and J. Yamagishi, "Articulatory control of HMM-based parametric speech synthesis using feature-space-switched multiple regression," *IEEE Trans. Audio Speech Lang. Processing*, vol. 21, no. 1, pp. 207–219, 2013.

[50] K. Tokuda, H. Zen, and A. Black, "An HMM-based speech synthesis system applied to English," in *Proc. IEEE Speech Synthetic Workshop, 2002, CD-ROM Proc.*

[51] K. Yu, H. Zen, F. Mairesse, and S. Young, "Context adaptive training with factorized decision trees for HMM-based statistical parametric speech synthesis," *Speech Commun.*, vol. 53, no. 6, pp. 914–923, 2011.

[52] J. Odell, "The use of context in large vocabulary speech recognition," Ph.D. thesis, Cambridge Univ., 1995.

[53] H. Zen, K. Tokuda, T. Masuko, T. Kobayashi, and T. Kitamura, "A hidden semi-Markov model-based speech synthesis system," *IEICE Trans. Inf. Syst.*, vol. E90-D, no. 5, pp. 825–834, 2007.

[54] T. Yoshimura, K. Tokuda, T. Masuko, T. Kobayashi, and T. Kitamura, "Duration modeling in HMM-based speech synthesis system," in *Proc. ICSLP*, 1998, vol. 2, pp. 29–32.

[55] M. Abe, S. Nakamura, K. Shikano, and H. Kuwabara, "Voice conversion through vector quantization," *J. Acoust. Soc. Jpn. (E)*, vol. 11, no. 2, pp. 71–76, 1990.

[56] Y. Stylianou, O. Cappe, and E. Moulines, "Continuous probabilistic transform for voice conversion," *IEEE Trans. Audio Speech Lang. Processing*, vol. 6, no. 2, pp. 131–142, 1998.

[57] D. Erro, A. Moreno, and A. Bonafonte, "Voice conversion based on weighted frequency warping," *IEEE Trans. Audio Speech Lang. Processing*, vol. 18, no. 5, pp. 922–931, 2010.

[58] S. Desai, A. W. Black, B. Yegnanarayana, and K. Prahallad, "Spectral mapping using artificial neural networks for voice conversion," *IEEE Trans. Audio Speech Lang. Processing*, vol. 18, no. 5, pp. 954–964, 2010.

[59] E. Helander, H. Silen, T. Virtanen, and M. Gabbouj, "Voice conversion using dynamic kernel partial least squares regression," *IEEE Trans. Audio Speech Lang. Processing*, vol. 20, no. 3, pp. 806C817, 2012.

[60] D. Saito, S. Watanabe, A. Nakamura, and N. Minematsu, "Statistical voice conversion based on noisy channel model," *IEEE Trans. Audio Speech Lang. Processing*, vol. 20, no. 6, pp. 1784–1794, 2012.

[61] K. Park and H. Kim, "Narrowband to wideband conversion of speech using GMM based transformation," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2000, pp. 1843–1846.

[62] A. Mouchtaris, J. Van der Spiegel, P. Mueller, and P. Tsakalides, "A spectral conversion approach to single-channel speech enhancement," *IEEE Trans. Audio Speech Lang. Processing*, vol. 15, no. 4, pp. 1180–1193, 2007.

[63] T. Toda, M. Nakagiri, and K. Shikano, "Statistical voice conversion techniques for body-conducted unvoiced speech enhancement," *IEEE Trans. Audio Speech Lang. Processing*, vol. 20, no. 9, pp. 2505–2517, 2012.

[64] T. Toda, A. Black, and K. Tokuda, "Statistical mapping between articulatory movements and acoustic spectrum using a Gaussian mixture model," *Speech Commun.*, vol. 50, pp. 215–227, 2008.

[65] H. Zen. (2013). Deep learning in speech synthesis. *Keynote speech given at ISCA SSW8.* [Online]. Available: http://research.google.com/pubs/archive/41539.pdf

[66] L. Deng and D. O'Shaughnessy, *Speech Processing: A Dynamic and Optimization-Oriented Approach.* New York: Marcel Dekker, 2003.

[67] J. Sun and L. Deng, "An overlapping-feature based phonological model incorporating linguistic constraints: Applications to speech recognition," *J. Acoust. Soc. Am.*, vol. 111, pp. 1086–1101, 2002.

[68] L. Deng, G. Ramsay, and D. Sun, "Production models as a structural basis for automatic speech recognition," *Speech Commun.*, vol. 33, nos. 2–3, pp. 93–111, Aug. 1997.

[69] L. Deng, "Switching dynamic system models for speech articulation and acoustics," in *Mathematical Foundations of Speech and Language Processing.* New York: Springer-Verlag, 2003, pp. 115–134.

[70] D. Yu, L. Deng, and G. Dahl, "Roles of pre-training and fine-tuning in context-dependent DBN-HMMs for real-world speech recognition," in *Proc. NIPS Workshop on Deep Learning and Unsupervised Feature Learning*, 2010.

[71] G. Dahl, D. Yu, L. Deng, and A. Acero, "Large vocabulary continuous speech recognition with context-dependent DBN-HMMs," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 4688–4691.

[72] G. Dahl, D. Yu, L. Deng, and A. Acero, "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition," *IEEE Trans. Speech Audio Processing*, vol. 20, no. 1, pp. 30–42, 2012.

[73] A. Mohamed, G. Dahl, and G. Hinton, "Acoustic modeling using deep belief networks," *IEEE Trans. Audio Speech Lang. Processing*, vol. 20, no. 1, pp. 14–22, 2012.

[74] T.N. Sainath, B. Kingsbury, H. Soltau, and B. Ramabhadran, "Optimization techniques to improve training speed of deep neural networks for large speech tasks," *IEEE Trans. Audio Speech Lang. Processing*, vol. 21, no. 11, pp. 2267–2276, 2013.

[75] X.-L. Zhang and Ji Wu, "Deep belief networks based voice activity detection," *IEEE Trans. Audio Speech Lang. Processing*, vol. 21, no. 4, pp. 697–710, 2013.

[76] B. Uria, S. Renals, and K. Richmond, "A deep neural network for acoustic-articulatory speech inversion," in *Proc. NIPS 2011 Workshop on Deep Learning and Unsupervised Feature Learning*, 2011.

[77] G. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural Computat.*, vol. 14, no. 8, pp. 1711–1800, 2002.

[78] G. Hinton, "Products of experts," in *Proc. 9th Int. Conf. Artificial Neural Networks*, 1999, pp. 1–6.

[79] G. Taylor, G. Hinton, and S. Roweis, "Modeling human motion using binary latent variables," in *Proc. Advances in Neural Information Processing Systems,* 2007, pp. 1345–1352.

[80] R. Neal, "Connectionist learning of belief networks," *Artificial Intell.*, vol. 56, no. 1, pp. 71–113, 1992.

[81] D. Rumelhart, G. Hinton, and R. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[82] S. Hochreiter, Y. Bengio, P. Frasconi, and J. Schmidhuber, "Gradient flow in recurrent nets: The difficulty of learning long-term dependencies," in *A Field Guide to Dynamical Recurrent Neural Networks*, S. Kremer and J. Kolen, Eds. Piscataway, NJ: IEEE Press, 2001, pp. 237–244.

[83] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," in *Proc. Advances in Neural Information Processing Systems*, 2007, pp. 153–160.

[84] H. Kawahara, I. Masuda-Katsuse, and A. de Cheveigne, "Restructuring speech representations using pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds," *Speech Commun.*, vol. 27, nos. 3–4, pp. 187–207, 1999.

[85] T. Fukada, K. Tokuda, T. Kobayashi, and S. Imai, "An adaptive algorithm for mel-cepstral analysis of speech," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 1992, vol. 1, pp. 137–140.

[86] T. Yoshimura, "Simultaneous modeling of phonetic and prosodic parameters, and characteristic conversion for HMM-based text-to-speech systems," Ph.D. thesis, Nagoya Inst. of Tech., 2002.

[SP]

©ISTOCKPHOTO.COM/H-H5800

Visual Domain Adaptation

[A survey of recent advances]

[Vishal M. Patel, Raghuraman Gopalan, Ruonan Li, and Rama Chellappa]

I n pattern recognition and computer vision, one is often faced with scenarios where the training data used to learn a model have different distribution from the data on which the model is applied. Regardless of the cause, any distributional change that occurs after learning a classifier can degrade its performance at test time. Domain adaptation tries to mitigate this degradation. In this article, we provide a survey of domain adaptation methods for visual recognition. We discuss the merits and drawbacks of existing domain adaptation approaches and identify promising avenues for research in this rapidly evolving field.

Supervised learning techniques have made tremendous contributions to machine learning and computer vision leading to the development of robust algorithms that are applicable in practical scenarios. While these algorithms have significantly advanced the state of the art, their performance is often limited by the amount of labeled training data available. Labeling is expensive and time-consuming due to the great amount of human effort involved. However, collecting unlabeled visual data

is becoming considerably easier due to the availability of low-cost consumer and surveillance cameras, and large Internet databases such as Flickr and YouTube. These data often come from multiple sources and modalities. Thus, when designing a classification or retrieval algorithm using these heterogeneous data, one has to constantly deal with the changing distribution of these data samples. Examples of such cases include: recognizing objects under poor lighting conditions and poses while algorithms are trained on well-illuminated objects at frontal pose, detecting and segmenting an organ of interest from magnetic resonance imaging (MRI) images when available algorithms are instead optimized for computed tomography and X-ray images, recognizing and detecting human faces on infrared images while algorithms are optimized for color images, etc.

This challenge is commonly referred to as *covariate shift* [1] or *data set bias* [2], [3]. Any distributional change or domain shift that occurs after training can degrade the performance at test time. For instance, in the case of face recognition, to achieve useful performance in the wild, face representation and recognition methods must learn to adapt to distributions specific to each application domain shown in Figure 1. Domain adaptation tackles this problem by leveraging domain shift characteristics

[FIG1] (a) Unconstrained face images. (b) Images with expression variations. (c) Images with pose variations. (d) Sketch images. Real-world object recognition algorithms, such as face recognition, must learn to adapt to distributions specific to each domain shown in (a)–(d) [91], [102], [103].

from labeled data in a related domain when learning a classifier for unseen data. Although some special kinds of domain adaptation problems have been studied under different names such as covariate shift [1], class imbalance [4], and sample selection bias [5], [6], it only started gaining significant interest very recently in computer vision. There are also some closely related but not equivalent machine-learning problems that have been studied extensively, including transfer learning or multitask learning [7], self-taught learning [8], semisupervised learning [9], and multiview analysis [10]. A review of domain adaptation methods from machine-learning and natural language processing communities can be found in [11]. Our goal in this article is to survey recent domain adaptation approaches for computer vision applications, discuss their advantages and disadvantages, and identify interesting open problems.

NOTATION AND RELATED LEARNING PROBLEMS
In this section, we introduce the notation and formulate the domain adaptation learning problem. Furthermore, we discuss the similarities and differences among the various learning problems related to domain adaptation.

NOTATION AND FORMULATION
We refer to the training data set with plenty of labeled data as the source domain and the test data set with a few labeled data or no labeled data as the target domain. Following [11], let X and Y denote the input (data) and the output (label) random variables, respectively. Let $P(X,Y)$ denote the joint probability distribution of X and Y. In domain adaptation, the target distribution is generally different than the source distribution and the true underlying joint distribution $P(X,Y)$ is unknown. We have two different distributions: one for the target domain and the other for the source domain. We denote the joint distribution in the source domain and

the target domain as $P_s(X,Y)$ and $P_t(X,Y)$, respectively. The marginal distributions of X and Y in the source and the target domains are denoted by $P_s(X)$, $P_s(Y)$, $P_t(X)$, $P_t(Y)$, respectively. Similarly, the conditional distributions in the two domains are denoted by $P_s(X|Y)$, $P_s(Y|X)$, $P_t(X|Y)$, $P_t(Y|X)$. The joint probability of $X=x$ and $Y=y$ is denoted by $P(X=x, Y=y) = P(x, y)$. Here, $x \in X$ and $y \in Y$, where X and Y denote the instance space and class label spaces, respectively.

Let $\mathcal{S} = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$, where $x^s \in \mathbb{R}^N$ denote the labeled data from the source domain. Here, x^s is referred to as an *observation*, and y^s is the *corresponding class label*. Labeled data from the target domain is denoted by $\mathcal{T}_l = \{(x_i^{tl}, y_i^{tl})\}_{i=1}^{N_{tl}}$, where $x^{tl} \in \mathbb{R}^M$. Similarly, unlabeled data in the target domain is denoted by $\mathcal{T}_u = \{x_i^{tu}\}_{i=1}^{N_{tu}}$, where $x^{tu} \in \mathbb{R}^M$. Unless specified otherwise, we assume $N = M$. Let $\mathcal{T} = \mathcal{T}_l \cup \mathcal{T}_u$. As a result, the total number of samples in the target domain is denoted by N_t, which is equal to $N_{tl} + N_{tu}$. Denote $\mathbf{S} = [x_1^s, \cdots, x_{N_s}^s]$ as the matrix of N_s data points from $\mathcal{S}$. Denote $\mathbf{T}_l = [x_1^{tl}, \cdots, x_{N_{tl}}^{tl}]$ as the matrix of N_{tl} data from $\mathcal{T}_l$, $\mathbf{T}_u = [x_1^{tu}, \cdots, x_{N_{tu}}^{tu}]$ as the matrix of N_{tu} data from $\mathcal{T}_u$ and $\mathbf{T} = [\mathbf{T}_l | \mathbf{T}_u] = [x_1^t, \cdots, x_{N_t}^t]$ as the matrix of N_t data from $\mathcal{T}$.

It is assumed that both the target and source data pertain to C classes or categories. Furthermore, it is assumed that all categories have some labeled data. We assume that there is always a relatively large amount of labeled data in the source domain and a small amount of labeled data in the target domain. As a result, $N_s \gg N_{tl}$.

The goal of domain adaptation is to learn a function $f(.)$ that predicts the class label of a novel test sample from the target domain. Depending on the availability of the source and target domain data, the domain adaptation problem can be defined in many different ways.

■ In semisupervised domain adaptation, the function $f(.)$ is learned using the knowledge in $\mathcal{S}$ and $\mathcal{T}_l$.

■ In unsupervised domain adaptation, the function $f(.)$ is learned using the knowledge in $\mathcal{S}$ and $\mathcal{T}_u$.

■ In multisource domain adaptation, $f(.)$ is learned from more than one domain in $\mathcal{S}$ accompanying each of the first two cases.

■ Finally, in the heterogeneous domain adaptation, the dimensions of features in the source and target domains are assumed to be different. In other words, $N \neq M$.

RELATED APPROACHES

COVARIATE SHIFT

One variation of the domain adaptation problem is where, given an observation, the conditional distributions of Y are the same in the source and the target domains, but the marginal distributions of X differ in the two domains. In other words, $P_t(Y|X=x) = P_s(Y|X=x)$ for all $x \in \mathcal{X}$, but $P_t(X) \neq P_s(X)$. This resulting difference between the two domains is known as *covariate shift* [1] or *sample selection bias* [5], [6].

Instance weighting methods can be used to address this covariate shift problem in which estimated weights are incorporated into a loss function in an attempt to make the weighted training distribution look like the testing distribution [11]. To see this, let us briefly review the empirical risk minimization framework for supervised learning [12]. Let $\theta \in \Theta$ be a model family from which we want to select an optimal parameter θ^* for the inference. Let $g(x, y, \theta)$ be a loss function. We want to minimize the following objective function:

$$\theta^* = \arg \min_{\theta \in \Theta} \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P(x, y) g(x, y, \theta)$$

to obtain the optimal θ^* for the distribution $P(X, Y)$. Since $P(X, Y)$ is unknown, we use the empirical distribution $\tilde{P}(X, Y)$ to estimate $P(X, Y)$. A good model $\hat{\theta}$ can be found by minimizing the following empirical risk:

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \tilde{P}(x, y) g(x, y, \theta)$$
$$= \arg \min_{\theta \in \Theta} \sum_{i=1}^{N} g(x_i, y_i, \theta),$$

where $\{(x_i, y_i)\}_{i=1}^{N}$ is a set of training instances randomly sampled from $P(X, Y)$. This formulation can be extended to domain adaptation by minimizing the following expected loss over the target domain distribution to find the optimal model parameter for the target domain [11]:

$$\theta_t^* = \arg \min_{\theta \in \Theta} \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} P_t(x, y) g(x, y, \theta).$$

In domain adaptation setting, the training instances $\{(x_i^s, y_i^s)\}_{i=1}^{N_s}$ are randomly sampled from the source distribution $P_s(X, Y)$. As a result, we get

$$\theta_t^* = \arg \min_{\theta \in \Theta} \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \frac{P_t(x, y)}{P_s(x, y)} P_s(x, y) g(x, y, \theta)$$
$$\approx \arg \min_{\theta \in \Theta} \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} \frac{P_t(x, y)}{P_s(x, y)} \tilde{P}_s(x, y) g(x, y, \theta)$$
$$= \arg \min_{\theta \in \Theta} \sum_{i=1}^{N_s} \frac{P_t(x_i^s, y_i^s)}{P_s(x_i^s, y_i^s)} g(x_i^s, y_i^s, \theta). \quad (1)$$

As can be seen from (1), weighting the loss of the source samples by $(P_t(x, y)/P_s(x, y))$ provides a solution to the domain adaptation problem [11].

Under covariate shift, the ratio $(P_t(x, y)/P_s(x, y))$ can be rewritten as:

$$\frac{P_t(x, y)}{P_s(x, y)} = \frac{P_t(x)}{P_s(x)} \frac{P_t(y|x)}{P_s(y|x)} = \frac{P_t(x)}{P_s(x)}.$$

As a result, one can weigh each training instance with $(P_t(x)/P_s(x))$. Shimodaira [1] explored this approach to reweight the log likelihood of each training instance using $(P_t(x)/P_s(x))$ for covariate shift. Various methods can be used to estimate the ratio $(P_t(x)/P_s(x))$. For instance, nonparametric density estimation [1], [13] and kernel mean match-based methods [14] have been proposed in the literature to directly estimate the ratio.

CLASS IMBALANCE

Another special case of the domain adaptation formulation assumes that $P_t(X|Y=y) = P_s(X|Y=y)$ for all $y \in \mathcal{Y}$, but $P_t(Y) \neq P_s(Y)$. This difference is often known as *class imbalance* [4]. Under this assumption, the ratio in (1) can be rewritten as:

$$\frac{P_t(x, y)}{P_s(x, y)} = \frac{P_t(y)}{P_s(y)} \frac{P_t(x|y)}{P_s(x|y)} = \frac{P_t(y)}{P_s(y)}.$$

As a result, one only needs to consider $(P_t(y)/P_s(y))$ to weigh the instances [15].

Resampling can also be applied on the training instances from the source domain so that the resampled data roughly has the same class distribution as the target domain. In these methods, underrepresented classes are oversampled and overrepresented classes are undersampled [11].

TRANSFER LEARNING

Multitask learning or transfer learning is closely related to domain adaptation [7], [16]. In multitask learning, different tasks are considered, but the marginal distribution of the source and target data are similar. In other words, assuming L tasks, the joint probability of each task $\{P(X, Y_i)\}_{i=1}^{L}$ is different, but there is only a single distribution $P(X)$ of the observation. When learning the class conditional models $\{P(Y_i|X, \theta_i)\}_{i=1}^{L}$ for L tasks, it is assumed that the model parameters of the individual tasks are drawn from a common prior distribution $P_\Theta(\theta)$.

> DOMAIN ADAPTATION IS A FUNDAMENTAL PROBLEM IN MACHINE LEARNING AND HAS GAINED A LOT OF TRACTION IN NATURAL LANGUAGE PROCESSING, STATISTICS, MACHINE LEARNING, AND, RECENTLY, IN COMPUTER VISION.

Since domain adaptation considers only a single task but different domains, it is a somewhat different problem than multitask learning. However, one can view domain adaptation as a special case of multitask learning with two tasks, one on the source domain and the other on the target domain. In fact, some domain adaptation methods are essentially solving transfer learning problems. We refer you to [16] for a comprehensive survey on various transfer learning methods.

SEMISUPERVISED LEARNING

The performance of a supervised classification algorithm is often dependent on the availability of a sufficient amount of training data. However, labeling samples is expensive and time-consuming due to the significant human effort involved. As a result, it is desirable to have methods that learn a classifier with high accuracy from only a limited amount of labeled training data. In semisupervised learning, unlabeled data are exploited to remedy the lack of labeled data. This in turn requires that the unlabeled data comes from the same distribution as the labeled data. Hence, if we ignore the domain difference, and treat the labeled source instances as labeled data and the unlabeled target domain instances as unlabeled data, then the resulting problem is that of the semisupervised learning problem. As a result, one can apply any semisupervised learning algorithm [9] to the domain adaptation problem. The subtle difference between domain adaptation and semisupervised learning comes from the following two facts [11]:

■ The amount of labeled data in semisupervised learning is small but large in domain adaptation.

■ The labeled data may be noisy in domain adaptation if one does not assume $P_s(Y|X = x) = P_t(Y|X = x)$ for all x, whereas, in semisupervised learning, the labeled data are assumed to be reliable.

In fact, there have been several works in the literature that extend semisupervised learning methods to domain adaptation. A naive Bayes' transfer classifier algorithm, which allows for the training and test data distributions to be different for text classification, was proposed in [17]. This algorithm first estimates the initial probabilities under a distribution of one labeled data set and then uses an expectation maximization (EM) algorithm to revise the model for a different distribution of the test data which are assumed to be unlabeled. This EM-based domain adaptation method can be shown to be equivalent to a semisupervised EM algorithm [18]. Some of the other methods that extend domain adaptation using semisupervised learning include [19] and [20].

SELF-TAUGHT LEARNING

Another problem related to domain adaptation and semisupervised learning is self-taught learning [8], [21]. In self-taught learning, we are given limited data for a classification task and also large amounts of unlabeled data that are only mildly related to the task. In particular, the unlabeled data may not arise from the same distribution or share the class labels. This assumption essentially differentiates self-taught learning from semisupervised learning. Self-taught learning is motivated by the observation that many randomly downloaded images contain basic

visual features, such as edges and corners, that are similar to those in the training images. As a result, if one is able to learn to recognize such patterns from the unlabeled data, then these features can be used for the supervised learning task of interest [8].

A sparse coding-based approach was proposed in [8] for self-taught learning, where a dictionary is learned using unlabeled data. Then, higher-level features are computed by solving a convex ℓ_1-regularized least squares problem using the learned dictionary and the labeled training data. Finally, a classifier is trained by applying a supervised learning algorithm such as a support vector machine (SVM) on these higher-level labeled features. A discriminative version of this algorithm was also presented in [22]. Furthermore, an unsupervised self-taught learning algorithm called *self-taught clustering* was proposed in [23]. Self-taught clustering aims at clustering a small collection of target unlabeled data with the help of a large amount of auxiliary unlabeled data. It is assumed that the target and auxiliary data have a different distribution. It was shown that this algorithm can greatly outperform several state-of-the-art clustering methods when using irrelevant unlabeled data.

MULTIVIEW ANALYSIS

In many computer vision applications, data often come in multiple views or styles. For instance, in object recognition, one has to deal with objects in different poses (views) and lighting conditions. As a result, one is faced with the problem of classifying or retrieving objects where the source (gallery) and target (query) data belong to different views. A direct comparison of instances across different views is not meaningful since they lie in different feature spaces.

In a multiview (also known as *cross-view* or *multimodal*) learning setting, correspondences are assumed to be known between the two view samples. In other words, samples are often given in pairs corresponding to different views. This assumption essentially differentiates cross-view learning from domain adaptation, where no correspondences are assumed between the domain samples. One popular solution in multiview learning is to learn view-specific projection directions using the paired samples from different views (domains) into a common latent space [10]. Classification or retrieval can then be performed in the latent space, where both the target and source data share the same feature space. Other methods for multiview learning include [24]–[28].

VISUAL DOMAIN ADAPTATION APPROACHES

Domain adaptation is a fundamental problem in machine learning and has gained a lot of traction in natural language processing, statistics, machine learning, and, recently, in computer vision. Early visual domain adaptation methods were applied to domain shift in videos [29], [30]. In particular, Duan et al. [30] proposed to adapt video concept classifiers between news videos collected from different news channels. Since then, there have been a plethora of approaches proposed in the vision literature for object category adaptation. In what follows, we present a number of recent domain adaptation strategies for visual recognition.

FEATURE AUGMENTATION-BASED APPROACHES

One of the simplest domain adaptation approaches is the feature augmentation work of Daumé III [31]. The goal is to make a domain-specific copy of the original features for each domain. Each feature in the original domain of dimension N is mapped onto an augmented space of dimension $3N$ simply by duplicating the feature vectors. The augmented feature maps for the source and target domains are defined as

$$\Phi^s(\mathrm{x}_i^s) = \begin{bmatrix} \mathrm{x}_i^s \\ \mathrm{x}_i^s \\ 0_N \end{bmatrix}, \qquad \Phi^t(\mathrm{x}_i^{tl}) = \begin{bmatrix} \mathrm{x}_i^{tl} \\ 0_N \\ \mathrm{x}_i^{tl} \end{bmatrix}, \qquad (2)$$

where $\mathrm{x}_i^s \in \mathcal{S}$, $\mathrm{x}_i^{tl} \in \mathcal{T}_l$, and 0_N denotes a zero vector of dimension N. The first N-dimensional component of this augmented feature corresponds to commonality between source and target, the second N-dimensional component corresponds to the source, while the last component corresponds to the target domain. Both source and target domain features are transformed using these augmented feature maps, and the resulting feature is passed onto the underlying supervised classifier. It was shown in [31] that when linear classifiers are used, this feature augmentation method is equivalent to decomposing the model parameter θ_i for domain i into $\tilde{\theta}_i + \theta_c$, where θ_c is shared by all domains. This "frustratingly easy" feature augmentation framework can be easily extended to a multidomain case by making more copies of the original feature space. Furthermore, a kernel version of this method is also derived in [31].

A feature augmentation-based method for utilizing the heterogeneous data from the source and target domains was recently proposed in [32]. The approach taken in [32] is to introduce a common subspace for the source and target data so that the heterogeneous features from two domains can be compared. In particular, both the source and target data of dimension N and M, respectively, are projected onto a latent domain of dimension l using two projection matrices $W_1 \in \mathbb{R}^{l \times N}$ and $W_2 \in \mathbb{R}^{l \times M}$, respectively. The augmented feature maps for the source and target domains in the common space are then defined as

$$\Phi^s(\mathrm{x}_i^s) = \begin{bmatrix} W_1\mathrm{x}_i^s \\ \mathrm{x}_i^s \\ 0_M \end{bmatrix} \in \mathbb{R}^{l+N+M}, \qquad (3)$$

$$\Phi^t(\mathrm{x}_i^{tl}) = \begin{bmatrix} W_2\mathrm{x}_i^{tl} \\ 0_N \\ \mathrm{x}_i^{tl} \end{bmatrix} \in \mathbb{R}^{l+N+M}, \qquad (4)$$

where $\mathrm{x}_i^s \in \mathcal{S}$, $\mathrm{x}_i^{tl} \in \mathcal{T}_l$, and 0_M is an M-dimensional zero vector. Once the data from both domains are transformed onto a common space, they can be readily passed onto a supervised classifier [32]. Figure 2 illustrates an overview of this method.

The general idea behind the frustratingly easy feature augmentation method of Daumé III [31] has been extended to consider a manifold of intermediate domains [33], [34]. Manifold-based methods for unsupervised visual domain adaptation were first proposed by Gopalan et al. [33]. Rather than working with the information conveyed by the source and target domains alone, [33] proposes using incremental learning by gradually following the geodesic path between the source and target domains. Geodesic flows are used to derive intermediate subspaces that interpolate between the source and target domains. Figure 3 shows an overview of this method.

It is assumed that the dimension of features in both the source and target domains is the same, e.g., $N = M$. First, principal component analysis (PCA) is applied on S and T_u, which generates two l-dimensional subspaces dented by two matrices S_1 and S_2, respectively, where $l < N$. The space of l-dimensional subspaces in $\mathbb{R}^N$ containing origin can be identified with the Grassmann manifold $\mathbb{G}_{N,l}$. As a result, S_1 and S_2 can be viewed as points on $\mathbb{G}_{N,l}$. By viewing $\mathbb{G}_{N,l}$ as quotient space of $SO(N)$, [here, $SO(N)$ represents the special orthogonal group, which is the group of orthogonal $N \times N$ matrices with determinant 1], the geodesic path in $\mathbb{G}_{N,l}$ starting from S_1 is given by a one-parameter exponential flow $\Psi(t') = Q \exp(t'B)J$, where exp refers to the matrix exponential, $Q \in SO(N)$ such that $Q^T S_1 = J$ and

$$J = \begin{bmatrix} I_l \\ 0_{N-l,l} \end{bmatrix}.$$

> **GEODESIC FLOWS ARE USED TO DERIVE INTERMEDIATE SUBSPACES THAT INTERPOLATE BETWEEN THE SOURCE AND TARGET DOMAINS.**
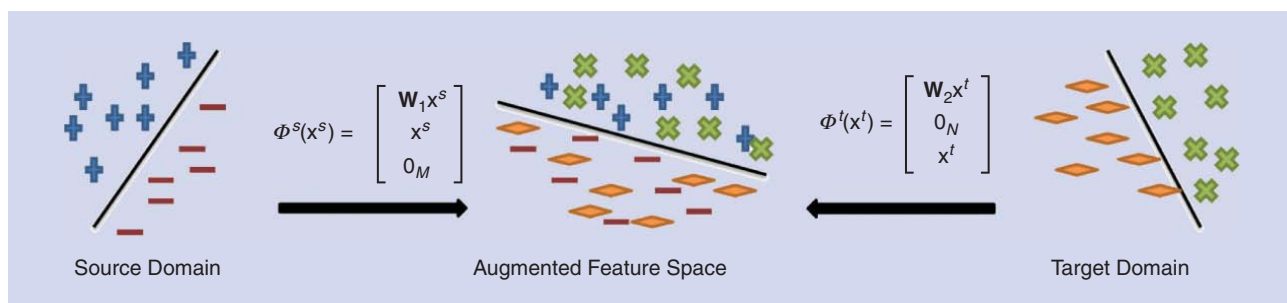
[FIG2] By using two projection matrices W_1 and W_2, one can transform the heterogeneous samples from two domains into an augmented feature space [32].

Source Domain $\qquad$ Augmented Feature Space $\qquad$ Target Domain

$$\Phi^s(\mathrm{x}^s) = \begin{bmatrix} W_1\mathrm{x}^s \\ \mathrm{x}^s \\ 0_M \end{bmatrix} \qquad \Phi^t(\mathrm{x}^t) = \begin{bmatrix} W_2\mathrm{x}^t \\ 0_N \\ \mathrm{x}^t \end{bmatrix}$$
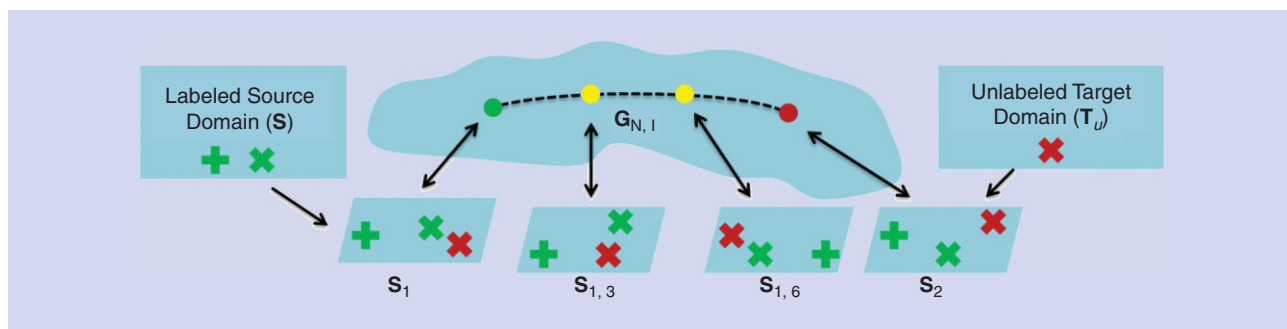
[FIG3] An overview of the manifold-based unsupervised domain adaptation method [33]. With labeled data **S** from source domain corresponding to two classes + and ×, and unlabeled data $\mathbf{T}_u$ from target domain belonging to class ×, generative subspaces $\mathbf{S}_1$ and $\mathbf{S}_2$ are derived using PCA. Then, by viewing $\mathbf{S}_1$ and $\mathbf{S}_2$ as points on the Grassmann manifold $\mathbb{G}_{N,l}$ (green and red circles), points along the geodesic between them (dashed line) are sampled to obtain geometrically meaningful intermediate subspaces (yellow circles).

Here, $\mathbf{I}_l$ is a $l \times l$ identity matrix and $\mathbf{B}$ is a skew-symmetric, block-diagonal matrix of the form

$$\mathbf{B} = \begin{bmatrix} 0 & \mathbf{A}^T \\ -\mathbf{A} & 0 \end{bmatrix},$$

$\mathbf{A} \in \mathbb{R}^{(N-l) \times l}$, where $(.)^T$ denotes the transposition operation and the submatrix $\mathbf{A}$ specifies the direction and the speed of geodesic flow. The geodesic flow between $\mathbf{S}_1$ and $\mathbf{S}_2$ is obtained by computing the direction matrix $\mathbf{A}$ such that the geodesic along that direction, while starting from $\mathbf{S}_1$, reaches $\mathbf{S}_2$ in unit time. The matrix $\mathbf{A}$ is computed using the inverse exponential mapping. Once $\mathbf{A}$ is computed, the expression for $\Psi(t')$ is used to obtain the intermediate subspaces between $\mathbf{S}_1$ and $\mathbf{S}_2$ by varying the value of t' between 0 and 1.

Let $\mathbf{S}'$ be the collection of subspaces $\mathbf{S}_t, t \in \mathbb{R}, 1 \leq t \leq 2$, which includes $\mathbf{S}_1$ and $\mathbf{S}_2$ and all intermediate subspaces. Let k denote the total number of such subspaces. The intermediate cross-domain data representations $\mathbf{U}$ are obtained by projecting the source data $\mathbf{S}$ and the target data $\mathbf{T}_u$ onto $\mathbf{S}'$. The final feature representation of dimension lk is obtained by projecting data onto k different subspaces. A model on these extended features is learned using partial least squares (PLS), and the assignment of target labels is performed using the nearest neighbor method [33]. A nonlinear version of this method, as well as an extension to semisupervised domain adaptation, has also been presented in [34]. Furthermore, assuming that the domain to which samples belong has been identified a priori [35], [36], this method has been extended to multidomain adaptation in [34].

Recently, the approach of [33] was kernelized and extended to the infinite case, defining a new kernel equivalent to integrating over all common subspaces that lie on the geodesic flow connecting the source and target subspaces $\mathbf{S}_1$ and $\mathbf{S}_2$, respectively [37]–[39]. Furthermore, assuming that the data lie in a union of subspaces in both the source and target domains, a framework based on the parallel transport of a union of the source subspaces on the Grassmann manifold was proposed in [40]. It was shown that this way of modeling data with a union of subspaces instead of a single subspace significantly improves the recognition performance [40].

FEATURE TRANSFORMATION-BASED APPROACHES

One of the earliest object category adaptation methods was proposed by Saenko et al. [41]. The idea behind this method is to adapt features across general image domains by learning transformations. Given feature vectors $\mathbf{x}^s \in \mathcal{S}$ and $\mathbf{x}^t \in \mathcal{T}$, a linear transformation $\mathbf{W} \in \mathbb{R}^{N \times M}$ from $\mathcal{T}$ to $\mathcal{S}$ is learned. The inner product similarity function between $\mathbf{x}^s$ and the transformed $\mathbf{x}^t$ is denoted by

$$\text{sim}_W = (\mathbf{x}^s)^T \mathbf{W} \mathbf{x}^t. \tag{5}$$

One can view this function as an inner product between the transformed target point $\mathbf{W}\mathbf{x}^t$ and $\mathbf{x}^s$. The objective is to learn the linear transformation given some form of supervision and then to use the learned similarity function in a classification algorithm [41]. A regularization function for the matrix $\mathbf{W}$ is introduced to avoid overfitting, which is denoted as $r(\mathbf{W})$. Assume that the supervision is a function of the learned similarity values sim_W, so a general optimization problem would seek to minimize the regularizer subject to supervision constraints given by functions c_i

$$\min_W r(\mathbf{W}) \text{ s.t. } c_i(\mathbf{S}^T \mathbf{W} \mathbf{T}) \geq 0, \quad 1 \leq i \leq J. \tag{6}$$

Equation (6) can be written as an unconstrained problem

$$\min_W r(\mathbf{W}) + \lambda \sum_i c_i(\mathbf{S}^T \mathbf{W} \mathbf{T}). \tag{7}$$

The regularizer studied in [41] is

$$r(\mathbf{W}) = \text{trace}(\mathbf{W}) - \log\det(\mathbf{W}), \tag{8}$$

and the resulting optimization problem is solved using an information-theoretic metric-learning [42] type of algorithm. One of the limitations of this method is that it can only be applied when the dimensionalities of the two domains are the same (e.g., $N = M$).

This work was extended in [43] by Kulis et al. to the more general case where the domains are not restricted to be the same dimensionality and arbitrary asymmetric transformations can be learned. Their method can deal with more general types of domain shifts and changes in feature type and dimension. Furthermore,

they show that the method in [41] is a special case of their general formulation, producing symmetric positive definite transformations [43]. It was shown that asymmetric indefinite transformations are more flexible for a variety of adaptation tasks than the symmetric transformations.

Recently, a low-rank approximation-based approach for semi-supervised domain adaptation was proposed in [44]. The basic goal of this method is to map the source data by a matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$ to an intermediate representation where each transformed sample can be reconstructed by a linear combination of the target data samples

$$\mathbf{WS} = \mathbf{T}_l \mathbf{Z}, \qquad (9)$$

where $\mathbf{Z} \in \mathbb{R}^{N_{tl} \times N_s}$ is the coefficient matrix. The following formulation is proposed to solve for the low-rank solution:

$$(\hat{\mathbf{W}}, \hat{\mathbf{Z}}, \hat{\mathbf{E}}) = \min_{\mathbf{W}, \mathbf{Z}, \mathbf{E}} \text{rank}(\mathbf{Z}) + \lambda \| \mathbf{E} \|_{2,1},$$
$$\text{s.t.} \quad \mathbf{WS} = \mathbf{T}_l \mathbf{Z} + \mathbf{E}, \quad \mathbf{WW}^T = \mathbf{I}, \qquad (10)$$

where $\text{rank}(.)$ denotes the rank of a matrix, λ is a parameter, $\mathbf{E} \in \mathbb{R}^{N \times N_s}$ is the error term, and the $\ell_{2,1}$-norm is defined as $\| \mathbf{E} \|_{2,1} = \sum_{i=1}^{N} \sqrt{\sum_{j=1}^{N_s} E_{ij}^2}$. As a common practice in rank minimization problems, the rank of $\mathbf{Z}$ is replaced by its nuclear norm in (10) [44]. The augmented Lagrange multiplier method is proposed to solve the optimization problem.

Once the solution $(\hat{\mathbf{W}}, \hat{\mathbf{Z}}, \hat{\mathbf{E}})$ is obtained, the source data are transformed to the target domain as

$$\hat{\mathbf{W}}\mathbf{S} - \hat{\mathbf{E}}. \qquad (11)$$

The transformed source data are mixed with the target samples as the augmented training samples for training the classifiers. The trained classifier is then used to perform recognition on the unseen test samples in the target domain [44]. An extension of this method for the multiple source domain adaptation problem has also been proposed in [44]. Other recent transformation-based visual domain adaptation methods include [45] and [46].

PARAMETER ADAPTATION METHODS

Several algorithms have been proposed in the literature that investigate modifying the SVM algorithms for the domain adaptation problem. In particular, Yang et al. proposed an adaptive SVM (A-SVM) [29] method in which the source classifier $f_S(\mathbf{x})$ trained on the source data $\mathcal{S} = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$ is adapted to a new classifier $f_T(\mathbf{x})$ for the unseen target data $\mathcal{T}_u = \{x_i^{tu}\}_{i=1}^{N_{tu}}$. The decision function is formulated as

$$f_T(\mathbf{x}) = f_S(\mathbf{x}) + \delta f(\mathbf{x}), \qquad (12)$$

where $\delta f(\mathbf{x})$ is the perturbation function. It was shown in [29] that the perturbation function can be formulated as $\delta f(\mathbf{x}) = \theta^T \phi(\mathbf{x})$, where a feature map ϕ is used to project $\mathbf{x}$ into a high-dimensional feature vector $\phi(\mathbf{x})$. The perturbation function $\delta f(\mathbf{x})$ is

learned using the labeled data $\mathcal{T}_l = \{(x_i^{tl}, y_i^{tl})\}_{i=1}^{N_{tl}}$ from the target domain. To learn the parameter $\mathbf{w}$ of the perturbation function $\delta f(\mathbf{x})$, the following optimization problem is solved:

$$\min_{\theta} \frac{1}{2} \| \theta \|^2 + \alpha \sum_{i=1}^{N_{tl}} \xi_i$$
$$\text{s.t.} \ \xi_i \geq 0,$$
$$y_i^{tl} f_S(x_i^{tl}) + y_i^{tl} \theta^T \phi(x_i^{tl}) \geq 1 - \xi_i, \forall (x_i^{tl}, y_i^{tl}) \in \mathcal{T}_l, \qquad (13)$$

where ξ_i is the penalizing variable and α is a parameter that determines how much error an SVM can tolerate. The first term in (13) tries to minimize the deviation between the new decision boundary and the old one, and the second term controls the penalty of the classification error over the training data in the target domain.

This work was improved in [47] for object category detection and in [48] for visual concept classification. Domain transfer SVM [49] attempts to reduce the mismatch in the domain distributions, measured by the maximum mean discrepancy (MMD) while also learning a target decision function. Other SVM-based domain adaptation methods include [50]–[54].

As discussed previously, several domain adaptation methods make use of the kernel methods. The classification performance of these kernel-based methods is highly dependent on the choice of the kernel. Multiple kernel learning (MKL) can be used to combine multiple kernel functions to obtained a better solution [55]. MKL has been shown to work well in many computer vision applications. However, these methods assume that both training and test data come from the same domain. As a result, MKL methods cannot learn the optimal kernel with the combined data from the source and target domains for the domain adaptation problem. Hence, training data from the auxiliary domain may degrade the performance of MKL algorithms in the target domain. To deal with this, several cross-domain kernel learning methods have been proposed in the literature [56]–[58].

In [56], adaptive MKL is used to learn a kernel function based on multiple base kernels. In [57], a kernel function and a classifier are simultaneously learned by minimizing both the structural risk functional and the distribution mismatch between the labeled and unlabeled samples from the auxiliary and target domains. It was shown in [56] and [57] that these domain-adaptive MKL methods can significantly outperform traditional MKL and cross-domain learning methods.

There are some limitations of the feature-based and parameter transfer-based visual domain adaptation methods reviewed in this survey. For instance, the transform-based approaches discussed in [41], [43], [45], and [46] are based on some notion of closeness between the transformed source samples and target samples. They do not optimize the objective function of a discriminative classifier directly. Also, the computational complexity of these methods is highly dependent on the total number of samples used for training. On the other hand, parameter adaptation-based methods such as [29] and [48] optimize the classifier directly but they are not able to transfer the adapted function to novel categories. To deal with this problem, several methods have been developed in the

literature that attempt to optimize both the transformation and classifier parameters jointly [59]–[61].

In particular, the max-margin domain transfer method was recently proposed by Hoffman et al. in [60], which uses an asymmetric transform $\mathbf{W}$ to map target features to a new representation where they are maximally aligned with the source and learns the transform jointly on all categories for which target labels are available. It provides a way to adapt max-margin classifiers in a multiclass setting by learning a common component of the domain shift as captured by $\mathbf{W}$.

The goal of this method is to jointly learn affine hyperplanes that separate the classes in the source domain and a transformation from the points in the target domain into the source domain such that the transformed target data lie on the correct side of the learned source hyperplanes. For simplicity, let us consider the optimization for the binary problem [60]

$$\min_{\mathbf{W},\theta,b} \frac{1}{2}\|\mathbf{W}\|_F^2 + \frac{1}{2}\|\theta\|_F^2$$

$$\text{s.t } y_i^s\left(\begin{bmatrix}\mathbf{x}_i^s\\1\end{bmatrix}^T\begin{bmatrix}\theta\\b\end{bmatrix}\right) \geq 1 \forall i \in \{1,\cdots,N_s\}$$

$$y_i^{tl}\left(\begin{bmatrix}\mathbf{x}_i^{tl}\\1\end{bmatrix}^T\mathbf{W}^T\begin{bmatrix}\theta\\b\end{bmatrix}\right) \geq 1 \forall i \in \{1,\cdots,N_{tl}\}, \tag{14}$$

where θ denotes the normal of the affine hyperplane and b is the bias term. This formulation can be easily extended to the multiclass case by adding a sum over the regularizers on all class-specifics parameters and adding the constraints for all categories. The resulting optimization problem is not convex. As a result, it is solved by alternating minimization on $\mathbf{W}$ and (θ, b) [60]. This work was extended in [61] to include Laplacian regularization using instance constraints that are encoded by an arbitrary graph.

Another approach to simultaneous learning of domain-invariant features and classifiers was proposed by Shi and Sha in [59]. Their framework is based on the notion of discriminative clustering in which both the source and target domains are assumed to be tightly clustered and clusters are assumed to correspond to class boundaries. It is assumed that for the same class, the clusters from the two domains are geometrically close to each other. Their formulation of learning the optimal feature space is based on maximizing the domain similarity that makes the source and the target domains look alike and minimizing the expected classification error on the target domain. An information-theoretic framework is proposed for solving their formulation [59].

DICTIONARY-BASED APPROACHES

The study of sparse representation of signals and images has attracted tremendous interest over the last few years. This is partly because signals or images of interest, although high dimensional, can often be coded using few representative atoms in some dictionary. In their seminal work, Olshausen and Field [62] introduced the idea of learning a dictionary from data instead of using off-the-shelf bases. Since then, data-driven dictionaries have been shown to work well for both image restoration and classification tasks [63], [64]. The efficiency of dictionaries in these wide range of applications can be attributed to the robust discriminant representations that they provide by adapting to particular data samples. However, the learned dictionary may not be optimal if the target data have a different distribution than the data used for training. Several dictionary-learning-based methods have been proposed in the literature to deal with this domain shift problem [65]–[68].

A function learning framework for the task of transforming a dictionary learned from one visual domain to the other while maintaining a domain-invariant sparse representation of a signal was proposed in [65]. Domain dictionaries are modeled by a linear or nonlinear parametric function. The dictionary function parameters and domain-invariant sparse codes are then jointly learned by solving an optimization problem. Motivated by the manifold-based incremental learning work of Gopalan et al. [33], [34], Ni et al. [67] proposed an unsupervised domain-adaptive dictionary-learning framework by generating a set of intermediate dictionaries, which smoothly connect the source and target domains. One of the important properties of this approach is that it allows the synthesis of data associated with the intermediate domains while exploiting the discriminative power of generative dictionaries. The intermediate data can then be used to build a classifier for recognition under domain shifts.

In [66], Shekhar et al. proposed a semisupervised domain-adaptive dictionary-learning framework for learning a single dictionary to optimally represent both source and target data. As the features may not be correlated well in the original space, they propose to project data from both the domains onto a common low-dimensional space while maintaining the manifold structure of the data. They argue that learning the dictionary on a low-dimensional space makes the algorithm faster and that irrelevant information in the original features can be discarded. Moreover, joint learning of dictionary and projections ensures that the common internal structure of data in both domains is extracted, which can be represented well by sparse linear combinations of dictionary atoms. Figure 4 shows an overview of this method [66].
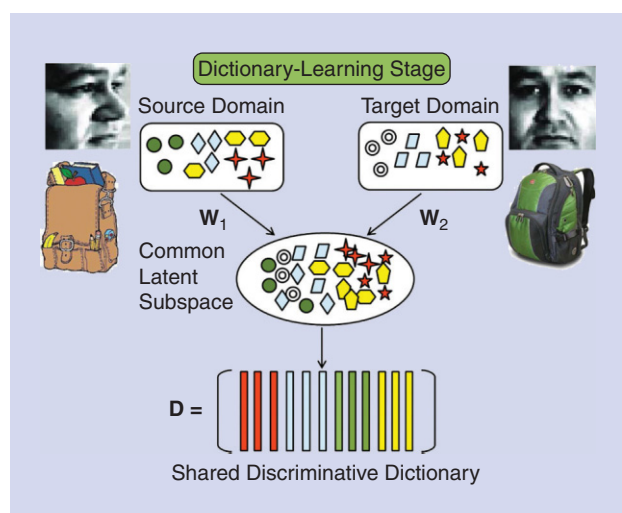
[FIG4] An overview of the domain-adaptive latent space dictionary-learning framework [66].

Given source and target domain data $\mathbf{S} \in \mathbb{R}^{N \times N_s}$ and $\mathbf{T}_l \in \mathbb{R}^{M \times N_{tl}}$, respectively, Shekhar et al. learn a shared K atom dictionary, $\mathbf{D} \in \mathbb{R}^{l \times K}$, and mappings $\mathbf{W}_1 \in \mathbb{R}^{l \times N}$ and $\mathbf{W}_2 \in \mathbb{R}^{l \times M}$ onto a common low-dimensional space, which will minimize the representation error in the projected space. Formally, the following cost is minimized:

$$C_1(\mathbf{D}, \mathbf{W}_1, \mathbf{W}_2, \mathbf{X}_1, \mathbf{X}_2) = \| \mathbf{W}_1 \mathbf{S} - \mathbf{D}\mathbf{X}_1 \|_F^2 + \| \mathbf{W}_2 \mathbf{T}_l - \mathbf{D}\mathbf{X}_2 \|_F^2$$

subject to sparsity constraints on $\mathbf{X}_1 \in \mathbb{R}^{K \times N_s}$ and $\mathbf{X}_2 \in \mathbb{R}^{K \times N_{tl}}$. It is assumed that rows of the projection matrices, $\mathbf{W}_1$ and $\mathbf{W}_2$, are orthogonal and normalized to unit norm. This prevents the solution from becoming degenerate, leads to an efficient scheme for optimization, and makes the kernelization of the algorithm possible. Note that this method does not require the data to be of the same dimension in the source and target domains. As a result, this method is applicable to heterogeneous domain adaptation problems [32].

To make sure that the projections do not lose too much information available in the original domains after projecting onto the latent space, a PCA-like regularization term is added, which preserves energy in the original signal, given as

$$C_2(\mathbf{W}_1, \mathbf{W}_2) = \| \mathbf{S} - \mathbf{W}_1^T \mathbf{W}_1 \mathbf{S} \|_F^2 + \| \mathbf{T}_l - \mathbf{W}_2^T \mathbf{W}_2 \mathbf{T}_l \|_F^2.$$

It is easy to show that the costs C_1 and C_2, after ignoring the constant terms in $\mathbf{Y}$, can be written as

$$C_1(\mathbf{D}, \tilde{\mathbf{W}}, \tilde{\mathbf{X}}) = \| \tilde{\mathbf{W}}\tilde{\mathbf{Y}} - \mathbf{D}\tilde{\mathbf{X}} \|_F^2, \tag{15}$$

$$C_2(\tilde{\mathbf{W}}) = -\text{trace}((\tilde{\mathbf{W}}\tilde{\mathbf{Y}})(\tilde{\mathbf{W}}\tilde{\mathbf{Y}})^T), \tag{16}$$

where

$$\tilde{\mathbf{W}} = [\mathbf{W}_1\ \mathbf{W}_2], \tilde{\mathbf{Y}} = \begin{pmatrix} \mathbf{S} & 0 \\ 0 & \mathbf{T}_l \end{pmatrix}, \text{and } \tilde{\mathbf{X}} = [\mathbf{X}_1\ \mathbf{X}_2].$$

Hence, the overall optimization is given as

$$\{\mathbf{D}^*, \tilde{\mathbf{W}}^*, \tilde{\mathbf{X}}^*\} = \arg \min_{\mathbf{D}, \tilde{\mathbf{W}}, \tilde{\mathbf{X}}} C_1(\mathbf{D}, \tilde{\mathbf{W}}, \tilde{\mathbf{X}}) + \lambda C_2(\tilde{\mathbf{W}})$$

$$\text{s.t. } \mathbf{W}_i \mathbf{W}_i^T = \mathbf{I}, i = 1, 2 \text{ and } \| \tilde{\mathbf{x}}_j \|_0 \leq T_0, \forall j, \tag{17}$$

where λ is a positive constant. An efficient two-step procedure is proposed for solving this optimization problem in [66]. Furthermore, this method has been extended to multiple domains and kernelized in [66]. Once the projection matrices and the dictionary are learned, given a novel test sample form the target domain, it is first projected onto the latent domain using $\mathbf{W}_2$ and classified using a variation of the latent sparse embedding residual classifier (LASERC) algorithm proposed in [69].

DOMAIN RESAMPLING
An unsupervised domain adaptation method was recently proposed in [70] and [71] based on the notion of landmarks. Landmarks are a subset of labeled data instances in the source domain that are distributed most similarly to the target domain [70].

The key insight of their method is that not all instances are created equally for adaptation. As a result, they pick out and exploit the most desirable instances to facilitate adaptation. An overview of this method is shown in Figure 5.

A variant of MMD is used to select samples from the source domain to match the distribution of the target domain. To identify landmarks, N_s indicator variables $\alpha = \{\alpha_i \in \{0,1\}\}$ are used, one for each data point in the source domain. If $\alpha_i = 1$, then $\mathbf{x}_i^s$ is regarded as a landmark. The vector α is identified by minimizing the MMD metric, defined with a kernel mapping function $\phi(\mathbf{x})$,

$$\min_\alpha \left\| \frac{1}{\sum_i \alpha_i} \sum_i \alpha_i \phi(\mathbf{x}_i^s) - \frac{1}{N_{tu}} \sum_j \phi(\mathbf{x}_j^{tu}) \right\|_{\mathcal{H}}^2$$

$$\text{s.t. } \frac{1}{\sum_i \alpha_i} \sum_i \alpha_i y_{ic} = \frac{1}{N_s} \sum_i y_{ic}, \tag{18}$$

where y_{ic} is the indicator variable for $y_{ic} = c$. The right-hand side of the constraint is simply the prior probability of the class c, estimated from the source domain.

The geodesic flow kernel computed between the source $\mathcal{S}$ and the target $\mathcal{T}_u$ is used to compose the kernel mapping function $\phi(\mathbf{x})$ [70]

$$\phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)$$
$$= \exp\{-(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{G}(\mathbf{x}_i - \mathbf{x}_j)/\sigma^2\}, \tag{19}$$

where $\mathbf{G}$ is computed using the singular value decomposition of $\mathbf{S}_1^T \mathbf{S}_2$. Here, $\mathbf{S}_1$ and $\mathbf{S}_2$ are the matrices obtained by applying PCA on $\mathbf{S}$ and $\mathbf{T}_u$, respectively [37].

A set of factors $\{\sigma_i \in [\sigma_{\min}, \sigma_{\max}]\}_{i=1}^Q$ is used to select the scale factor σ in (19). For each σ_i, (18) is solved to obtain the corresponding landmarks $\mathcal{L}^i$ whose α_i is equal to one. For each set of landmarks, a new domain pair is constructed by moving the landmarks from the original source to the target domains. It was argued that each auxiliary task is easier to adapt than the original pair $\mathcal{S}$ and $\mathcal{T}_u$ [70].
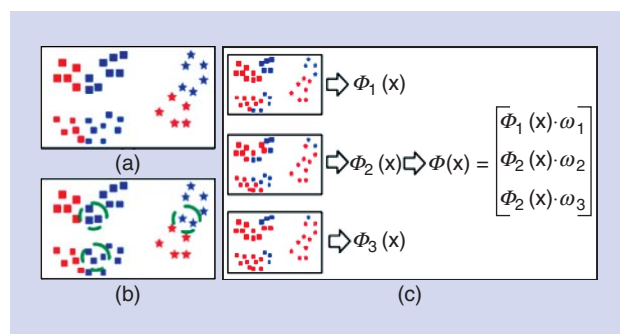
[FIG5] An overview of the landmark-based method proposed in [70]. (a) The original domain adaptation problem where the instances in red are from the target and those in blue are from the source. (b) Landmarks, shown inside the green circles, are data instances from the source that can be regarded as samples from the target. (c) Multiple auxiliary tasks are created by augmenting the original target with landmarks, which switches their color from blue to red. Each task gives rise to a new feature representation. These representations are combined discriminatively to form domain-invariant features for the original domain adaptation problem [70].
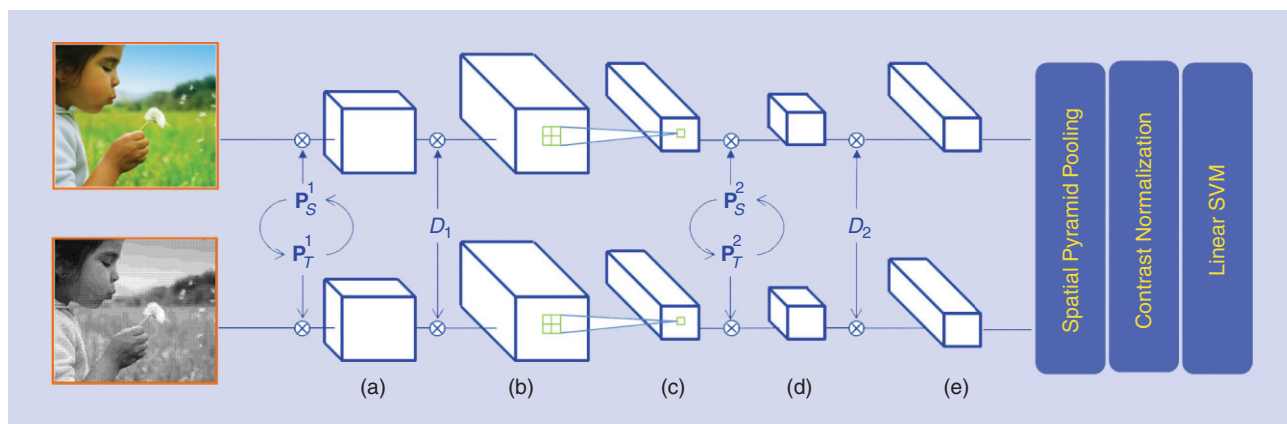
[FIG6] An illustration of domain adaptation using a sparse and hierarchical network (DASH-N) algorithm [77]. The source domain is RGB images, and the target domain is half-tone images. First, the images are divided into small overlapping patches. These patches are vectorized while maintaining their spatial arrangements. (a) Performing contrast normalization and dimensionality reduction using P_S for source images and P_T for target images. The circular feedbacks between P_S and P_T indicate that these two transformations are learned jointly. (b) Obtaining sparse codes using the common dictionary D_1. (c) Performing max pooling. The process then repeats for (d) and (e) layer 2, except that the input is the sparse codes from layer 1 instead of pixel intensities. At the final stage, spatial pyramids with max pooling are used to create image descriptors. Classification is done using a linear SVM.

The final kernel is then learned as a convex combination of all of the kernels from the auxiliary tasks

$$\mathbf{F} = \sum_i \beta_i \mathbf{G}_i \quad \text{s.t. } \beta_i \geq 0 \text{ and } \sum_i \beta_i = 1. \tag{20}$$

The coefficients β_i are optimized on a labeled training set $\sum_i \mathcal{L}^i$ composed of all landmarks selected at different granularities. Finally, $\mathbf{F}$ is used in an SVM classifier whose accuracy is optimized with the standard MKL algorithm to learn β_i [70], [71]. Since $\sum_i \mathcal{L}^i$ consists of landmarks that are distributed similarly to the target, it is expected that the classification error on $\sum_i \mathcal{L}^i$ will be a good proxy to that of the target domain [70].

OTHER METHODS

Deep neural networks have had tremendous success achieving a state-of-the-art performance on a number of machine-learning and computer vision tasks [72]. This is due in part to the fact that deep networks are able to learn extremely powerful hierarchical nonlinear representations of the inputs [73], [74]. Motivated by recent works on deep learning, several hierarchical domain adaptation approaches have been proposed in the literature [75]–[79].

In [78], multiple intermediate representations are explored along an interpolating path between the target and source domains. Starting with all the source data samples $\mathcal{S}$, intermediate sampled data sets are generated. For each successive data set, the proportion of samples randomly drawn from $\mathcal{T}$ is increased and the proportion of samples drawn from $\mathcal{S}$ is decreased. Let $i \in [1, \cdots, k]$ be an index set over k intermediate data sets. Then, $\mathcal{S}_i = \mathcal{S}$ for $i = 1$, $\mathcal{S}_i = \mathcal{T}$, for $i = k$. For $i \in [2, \cdots, k-1]$, data sets $\mathcal{S}_i$ and $\mathcal{S}_{i+1}$ are created in a way so that the proportion of samples from $\mathcal{T}$ in $\mathcal{S}_i$ is less than in $\mathcal{S}_{i+1}$. Each of these data sets can be thought of as a single point on a particular kind of interpolating path between $\mathcal{S}$ and $\mathcal{T}$.

For each intermediate data set $\mathcal{S}_i$, a deep nonlinear feature extractor is trained. Once feature extractors corresponding to all points on the path are trained, any input sample can be represented by concatenating all of the outputs from the feature extractors together to create path features for the input. The hope is that this path representation will be more effective at domain adaptation because it is constructed to capture information about incremental changes between the source and target domains similar to [33] and [37]. After creating the path representation of the inputs, a classifier is trained on the data generated from the source domain data by minimizing an appropriate loss function [78].

Another recent work for visual domain adaptation using hierarchical networks was recently proposed by Nguyen et al. [77]. Their method jointly learns a hierarchy of features together with transformations that address the mismatch between different domains. This method was motivated by [80] in which multilayer sparse coding networks are proposed for building feature hierarchies layer by layer using sparse codes and spatial pooling. Figure 6 shows an overview of the sparse hierarchical domain adaptation method [77]. The network contains multiple layers, each of which contains three sublayers. The first sublayer performs contrast normalization and dimensionality reduction on the input data. Sparse coding is carried out in the second sublayer. In the final sublayer, adjacent features are max-pooled together to produce a new feature. The output from one layer becomes the input to the next layer. This method can be viewed as a generalization of the domain-adaptive dictionary learning framework [66] using hierarchical networks. An extension of this method to multiple source domains has also been presented in [77].

Visual attributes are human understandable properties to describe images such as blue, dark, and two-legged. They are valuable as a semantic cue in various vision problems. Recent research

explores a variety of applications for visual attributes including face verification [81], object recognition [82]–[84], and facilitating transfer learning [85]. The existing methods [82], [84], [85] assume that one model of an attribute is sufficient to capture all user perceptions. However, there are some real perceptual differences between annotators. Consider the example shown in Figure 7: five users confidently declared that the shoe on the left is formal, while five confidently declared the opposite. These differences stem from several factors such as the words for attributes are imprecise, their meaning often depends on context and culture, and they often stretch to refer to quite distinct object categories [86].

To capture the inherent differences in perception, [86] proposes to model attributes in a user-specific way. In particular, attribute learning is posed as an adaptation problem. First, they leverage any commonalities in perception to learn a generic prediction function using a large margin learning algorithm and data labeled with a majority vote from multiple annotators. Then, they use a small number of user-labeled examples to adapt the parameters of the generic model into a user-specific prediction function while not straying too far from the prior generic model. Essentially, this amounts to imposing regularizers on the learning objective favoring user-specific model parameters that are similar to the generic ones while still satisfying the user-specific label constraints [86]. The impact of this attribute adaptation work is that one can capture a user's perception with minimal annotation effort. It was shown that the resulting personalization can make attribute-based image searches more accurate [86].

Tommasi and Caputo [87] very recently proposed a naive Bayes' nearest neighbor-based domain adaptation method that iteratively learns a Mahalanobis class-specific metric while inducing for each sample a large margin separation among classes. Both semisupervised and unsupervised domain adaptation scenarios are presented.

In [88], Jain and Farfade proposed an approach for adapting a cascade of classifiers to perform classification in a similar domain for which only a few positive examples are available. A cascade of classifiers is a classifier f that is composed of m stage classifiers $\{f_1, \cdots, f_m\}$ that are applied in a sequential manner. They are commonly used for anomaly detection and one-class classification. It was shown that, by adapting classification cascades to new domains, one can obtain huge gains in performance in detecting faces of human babies and human-like characters from movies.

APPLICATIONS

In this section, we illustrate through different application examples the uses and capabilities of various visual domain adaptation methods. In particular, we focus on object recognition and face recognition applications.

FACE RECOGNITION

Face recognition is a challenging problem that has been actively researched for more than two decades [89]. The current systems work very well when training, and test images are captured under controlled conditions. However, their performance degrades significantly when the test images contain variations that are not present in the training images. One of these variations is change in pose. Along with the frontal images with different illumination (source images), if we are also given a few images at different poses (target images), then the resulting face recognition problem can be viewed a domain adaptation problem [65], [66], [90].

Face recognition experiments were conducted on the Carnegie Mellon University (CMU) multipose, illumination, and expression (PIE) data set [91] with images of 129 subjects in a frontal pose as the source domain, and five other off-frontal poses as the target domain. Images under five illumination conditions across source and target domains were used for training with which images from the remaining 15 illumination conditions in the target domain were recognized. The results provided in Table 1 show that the dictionary-based adaptation method [66] compares favorably with some of the recently proposed multiview recognition algorithms [10] as well as many other nonadaptation techniques and gives the best performance on average. Note that the discriminative dictionary-learning algorithm, Fisher discrimination dictionary learning (FDDL) [92], does not provide the best results here as it is not able to efficiently represent the nonlinear changes introduced by the pose variation.

Furthermore, the learned dictionaries were also used for pose alignment where the goal is to align faces from one pose to a different pose. This is a challenging problem since actual pose variations

> **VISUAL ATTRIBUTES ARE HUMAN UNDERSTANDABLE PROPERTIES TO DESCRIBE IMAGES SUCH AS BLUE, DARK, AND TWO-LEGGED.**

[FIG7] Virtual attribute interpretations vary slightly from viewer to viewer. For instance, five viewers confidently declare the shoe as (a) formal or (b) more ornamented, while five others confidently declare the opposite. Attribute adaptation models are proposed to take these differences in perception into account [86].

[TABLE 1] A COMPARISON OF VARIOUS ALGORITHMS FOR FACE RECOGNITION ACROSS POSE [66].

METHOD	PROBE POSE					AVERAGE
	15°	30°	45°	60°	75°	
PCA	15.3	5.3	6.5	3.6	2.6	6.7
PLS [27]	39.3	40.5	41.6	41.1	38.7	40.2
LDA	98	94.2	91.7	84.9	79	89.5
CCA [27]	92.1	89.7	88	86.1	83	83.5
GMLDA [10]	**99.7**	**99.2**	98.6	94.9	95.4	97.6
FDDL [92]	96.8	90.6	94.4	91.4	90.5	92.7
SDDL [66]	98.4	98.2	**98.9**	**99.1**	**98.8**	**98.7**

Boldface indicates the top performing algorithm in each experiment.
LDA: Linear discriminant analysis; GMLDA: generalized multiview linear discriminant analysis; and SDDL: shared domain-adapted dictionary learning.
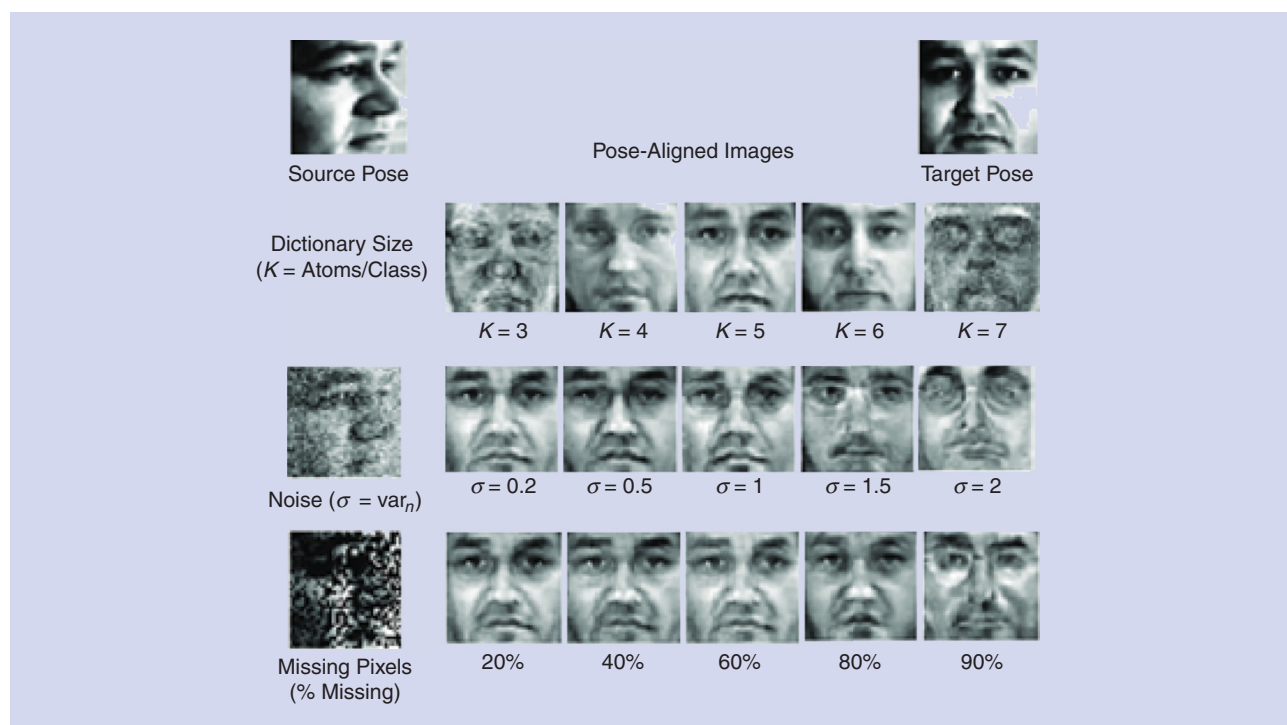
[FIG8] Examples of pose-aligned images. Synthesis in various conditions demonstrate the robustness of the domain-adaptive dictionary-learning method [66].

are three dimensional (3-D), whereas the image evidence one has is two dimensional (2-D). Sample results are shown in Figure 8. One of the interesting features of the dictionary-based adaptation methods is that they allow the synthesis of data associated with different domains while exploiting the generative power of dictionary-based representations. This is essentially what is highlighted in the last two rows of Figure 8. The dictionary-based method is robust at high levels of noise and missing pixels. It produces denoised and inpainted synthesized images. Additional results on various face recognition tasks using domain adaptation can be found in [65] and [67].

OBJECT RECOGNITION

In this section, we compare the performance of various visual domain adaptation methods on a benchmark object recognition data set that was introduced in [41]. The data set consists of images from three sources: Amazon (consumer images from online merchant sites), digital single-lens reflex [(DSLR) images from a DSLR camera], and Webcam (low-quality images from Webcams). In addition, algorithms are tested on the Caltech-256 data set [93], taking it as the fourth domain. Figure 9 shows sample images from these data sets and clearly highlights the differences between them.

Three setups are followed for comparing the performance of various algorithms. In the first setup, ten classes: "Backpack," "Touring Bike," "Calculator," "Headphones," "Computer Keyboard," "Laptop 101," "Computer Monitor," "Computer Mouse," "Coffee Mug," and "Video Projector" common to all the four data sets are used. In this case, there are a total of 2,533 images. Each category has eight to 151 images in a data set. In the second setup, all 31

classes from Amazon, Webcam, and DSLR are used to evaluate various algorithms. Finally, in the third setup, methods for adaptation are evaluated using multiple domains. In this case, the first data set is used, and the methods are tested on all 31 classes in it. For both cases, we use 20 training samples per class for Amazon/Caltech, eight samples per class for DSLR/Webcam when used as source, and three training samples for all of them when used for the target domain. The rest of the data in the target domain is used for testing. The experiment is run multiple times for random train/test splits, and the result is averaged over all the runs. For the unsupervised case, the same setting as semisupervised adaptation described earlier is followed but without using any labeled data from the target domain. (Several recent methods explore both source and target data at once in a transductive manner rather than splitting the data sets into multiple training/testing partitions; see [70] for details on the evaluation protocol using this setting.)

SEMISUPERVISED ADAPTATION
RESULTS USING A SINGLE SOURCE

The semisupervised adaptation recognition results of different algorithms on eight pairs of source–target domains and on all 31 classes are shown in Tables 2 and 3, respectively. The baseline results obtained using the hierarchical matching pursuit (HMP) method [80] as well as the FDDL method [92], which learn the dictionaries separately for the source and target domains without performing domain adaptation, are also included.

Compared to the metric-learning-based approach [41], manifold-based feature concatenation methods [33], [37] provide better results. This makes sense because, by finding intermediate domain

[FIG9] Some example images from the "Keyboard" and "Backpack" categories in Caltech-256, Amazon, Webcam, and DSLR. The Caltech-256 and Amazon data sets have diverse images, and Webcam and DSLR are similar data sets with images mostly from offices [66].

[TABLE 2] THE SEMISUPERVISED DOMAIN ADAPTATION RESULTS OF DIFFERENT APPROACHES ON FOUR DOMAINS WITH TEN COMMON CLASSES (C: CALTECH, A: AMAZON, D: DSLR, W: WEBCAM).

METHODS	C → A	C → D	A → C	A → W	W → C	W → A	D → A	D → W
METRIC [41]	33.7 ± 0.8	35 ± 1.1	27.3 ± 0.7	36 ± 1	21.7 ± 0.5	32.3 ± 0.8	30.3 ± 0.8	55.6 ± 0.7
SGF [33]	40.2 ± 0.7	36.6 ± 0.8	37.7 ± 0.5	37.9 ± 0.7	29.2 ± 0.7	38.2 ± 0.6	39.2 ± 0.7	69.5 ± 0.9
GFK [37]	46.1 ± 0.6	55 ± 0.9	39.6 ± 0.4	56.9 ± 1	32.8 ± 0.1	46.2 ± 0.6	46.2 ± 0.6	**80.2 ± 0.4**
FDDL [92]	39.3 ± 2.9	55 ± 2.8	24.3 ± 2.2	50.4 ± 3.5	22.9 ± 2.6	41.1 ± 2.6	36.7 ± 2.5	65.9 ± 4.9
HMP [80]	67.7 ± 2.3	70.2 ± 5.1	51.7 ± 4.3	70 ± 4.2	46.8 ± 2.1	61.5 ± 3.8	64.7 ± 2	76.0 ± 4
SDDL [66]	49.5 ± 2.6	76.7 ± 3.9	27.4 ± 2.4	72 ± 4.8	29.7 ± 1.9	49.4 ± 2.1	48.9 ± 3.8	72.6 ± 2.1
DASH-N [77]	**71.6 ± 2.2**	**81.4 ± 3.5**	**54.9 ± 1.8**	**75.5 ± 4.2**	**50.2 ± 3.3**	**70.4 ± 3.2**	**68.9 ± 2.9**	77.1 ± 2.8

Boldface indicates the top-performing algorithm in each experiment.
SGF: Subspaces by sampling geodesic flow; GFK: geodesic flow kernel.

representations, one is able to learn a feature vector that is more robust than a feature vector that results by learning a single transformation that minimizes the effect of the domain shift. The SDDL method can be viewed as an extension of the FDDL method, which simultaneously learns discriminative dictionaries on a latent space where both the source and the target data are forced to have similar sparse representation. As a result, one can clearly see the performance gain of the SDDL method over the FDDL method as well as the manifold-based methods in Tables 2 and 3.

The HMP method [80] builds a feature hierarchy layer by layer using an efficient matching pursuit encoder. It consists of three

[TABLE 3] SINGLE-SOURCE SEMISUPERVISED DOMAIN ADAPTATION RESULTS ON ALL 31 CLASSES.

METHOD	A → W	D → W	W → D
METRIC [41]	44	31	27
RDALR [44]	50.7 ± 0.8	36.9 ± 19.9	32.9 ± 1.2
SGF [33]	57 ± 3.5	36 ± 1.1	37 ± 2.3
GFK [37]	46.4 ± 0.5	61.3 ± 0.4	66.3 ± 0.4
HMP [80]	55.7 ± 2.5	50.5 ± 2.7	56.8 ± 2.6
SDDL [66]	50.1 ± 2.5	51.2 ± 2.1	50.6 ± 2.6
DASH-N [77]	**60.6 ± 3.5**	**67.9 ± 1.1**	**71.1 ± 1.7**

Boldface indicates the top performing algorithm in each experiment.
RDALR: Robust domain adaptation with low-rank reconstruction.

[TABLE 4] MULTIPLE-SOURCE DOMAIN ADAPTATION RESULTS OF VARIOUS METHODS ON THE AMAZON, WEBCAM, AND DSLR DATA SETS.

SOURCE	TARGET	SGF [34]	SGF [33]	RDALR [44]	FDDL [92]	SDDL [66]	A-SVM [29]	HMP [80]	DASH-N [77]
DSLR, AMAZON	WEBCAM	**64.5 ± 0.3**	52 ± 2.5	36.9 ± 1.1	41 ± 2.4	57.8 ± 2.4	30.4 ± 0.6	47.2 ± 1.9	64.5 ± 2.3
AMAZON, WEBCAM	DSLR	51.3 ± 0.7	39 ± 1.1	31.2 ± 1.3	38.4 ± 3.4	56.7 ± 2.3	25.3 ± 1.1	51.3 ± 1.4	**68.6 ± 3.7**
WEBCAM, DSLR	AMAZON	38.4 ± 1.0	28 ± 0.8	20.9 ± 0.9	19 ± 1.2	24.1 ± 1.6	17.3 ± 0.9	37.3 ± 1.4	**41.8 ± 1.1**

main components: batch tree orthogonal matching pursuit, spatial pyramid matching, and contrast normalization. As a result, it is robust to some of the variations present in the images such as illumination changes, pose variations, and resolution variations. The DASH-N method essentially extends the SDDL and HMP methods by learning features directly from data for domain adaptation. As a result, it provides a more robust and discriminative representation of the data and performs the best on this data set on both settings. The dictionary learning-based methods [92], [80] essentially find the common internal structure of the data. They inherently have the denoising capability and provide robust representation of the data. This is one of the reasons why in some cases the FDDL and the HMP methods provide better results than metric-learning- and manifold-based methods.

SEMISUPERVISED ADAPTATION RESULTS
USING MULTIPLE SOURCES
As some of the methods reviewed in this paper can also handle multiple domains, we report results of different algorithms on multiple-source adaption. Table 4 shows the results for three possible combinations. Again, the sparse hierarchical network-based adaptation method [77] performs the best. The incremental learning motivated manifold method [34] also provides good results on multidomain adaptation using this data set. It is interesting to see that increasing the number of domains can be helpful, especially when compared to a single source and single target. Many multidomain adaptation methods in Table 4 outperform a single source and a single target in many cases, although, in a small number of cases, they do not outperform a single source and a single target. As a result, a better strategy to deal with multiple domains is required in these cases.

UNSUPERVISED DOMAIN ADAPTATION RESULTS
The results of three source–target combinations of the Amazon, DSLR, and Webcam data sets are shown in Table 5. The manifold-based approach [34] outperforms the existing unsupervised domain adaptation methods in two of the three source–target combinations. The information–theoretic learning method [59] for unsupervised domain adaptation also performs well on this data set. By comparing the results in Tables 2 and 3 with the results in Table 5, we see that the semisupervised adaptation results are generally better than in the unsupervised case. Using labels in both the intermediate data generation and classification stage generally produces better results than using labels only during classification [34]. Also, it is interesting to see that, since the introduction of this data set in [41], the recognition performance has significantly improved in the last few years.

COMPUTATIONAL COMPLEXITY
The main processing steps involved in manifold-based adaptation techniques [33], [34], [37] are computing the geodesic between the source and target domains, and then sampling points along the geodesic to infer intermediate domains that account for the domain shift. This involves mapping entities on the manifold to the locally Euclidean tangent plane and warping the results from the tangent plane back onto the manifold. Computationally efficient algorithms for these steps have been discussed in the literature for Grassmann manifolds [94]. For orthogonal matrices of dimensions $N_1 \times N_2$, the geodesic computation has a complexity of $O(N_1^2 N_2)$ along with an $O(N_1 N_2)$ cost for sampling each point along the geodesic.

For deep learning approaches [77], [78], the complexity depends, among others, on the number of layers used in the hierarchy to learn feature correlation for adaptation. While the deep network circuits can have different architectures, such as autoencoders and restricted Boltzmann machines, there is an active stream of work in making the training procedure of these circuits computationally tractable. See [72] for a more detailed discussion on the complexity of deep architectures.

A major computationally heavy step of dictionary-based domain adaptation methods is dominated by sparse coding. Efficient batch methods have been proposed to learn dictionaries for large-scale problems. For instance, a batch orthogonal matching pursuit-based KSVD algorithm for learning dictionaries was proposed in [95]. It was shown that the operation count per training iteration for learning a dictionary of size $l \times K$ with R number of training signals are $R(T_0^2 K + 2lK)$, where T_0 is the target sparsity. One can also adapt fast ℓ_1 solvers for sparse coding [96], [97] rather than using greedy orthogonal matching pursuit algorithms.

For the low-rank approximation-based methods, the major computation is in finding the SVD of a matrix. As a result, these methods tend to be time-consuming if the matrix is large. However, efficient methods do exist for finding the low-rank approximation of large matrices [98]–[100].

[TABLE 5] UNSUPERVISED DOMAIN ADAPTATION RESULTS OF VARIOUS METHODS ON THE AMAZON, WEBCAM, AND DSLR DATA SETS.

SOURCE	TARGET	SGF [34]	SGF [33]	RDALR [44]	GFK [37]	ITLUDA [59]
WEBCAM	DSLR	**71.2**	19 ± 1.2	32.89 ± 1.2	49.7 ± 0.5	—
DSLR	WEBCAM	68.8	26 ± 0.8	36.85 ± 1.9	44.6 ± 0.3	**83.6 ± 0.5**
AMAZON	WEBCAM	**55.6**	39 ± 2	50.71 ± 0.8	15 ± 0.4	38.5 ± 1.3

Many parameter adaptation methods such as A-SVM [29] are large-scale quadratic programming problems for which efficient implementations do exist in the literature (see [101] for more details).

CONCLUSIONS AND FUTURE DIRECTIONS

This article attempted to provide an overview of recent developments in domain adaptation for computer vision, with an emphasis on applications to the problems of face and object recognition. We believe that the availability of massive data has brought substantial opportunities and challenges to the analysis of data sets bias or covariant shifts and domain adaptation problems. We hope that the survey has helped to guide interested readers among the extensive literature to some degree, but obviously it cannot cover all of the literature on domain adaptation, and we have chosen to focus on a representative subset of the latest progress made in computer vision.

Domain adaptation promises to be an active area of research, especially as one of the possible ways to quickly propagate semantic annotations to the large-scale visual data being acquired every minute. In computer vision, researchers have identified specific challenges that do not belong to machine learning: a major question among them that is rarely addressed in traditional domain adaptation research is one of adapting structured (nonvector) data representations. In machine learning or natural language processing, an input sample is usually represented as a vector in Euclidean space, different samples are treated as independent observations, and the task is typically classification. This is, however, not the case in computer vision where the representations to be potentially adapted include shapes and contours, deformable and articulated 2-D or 3-D objects, graphs, and random fields, intrinsic images, as well as visual dynamics, none of which is directly supported by vectorial domain adaptation techniques. In addition to recognition and detection, models and algorithms for segmentation, reconstruction, and tracking are awaiting mechanisms that do not yet exist to be adapted toward emerging new domains. All of these challenges necessitate continuous efforts on characterizing visual domain shift and a paradigm of effective and efficient adaptation methods that are dedicated to visual data.

In the meantime, it is generally accepted that domain shifts in computer vision are usually due to causes from the imaging process that can be explained physically, such as illumination changes, sensor changes, and viewpoint changes. We believe that incorporating these physical priors into strong statistical adaptation approaches will not only lead to a performance increase but also to other insights in understanding the imaging process. This calls for a physically informed adaptation paradigm that better exploits knowledge about image formation and better integrates other domain-specific knowledge implied by the diverse set of partial, noisy, and multimodal side information accompanying the visual data, such as imagery obtained from online social media. We hope that by appropriately incorporating a physically informed adaptation paradigm, distributional changes across different sensors (electro-optical/synthetic aperture radar, infrared/synthetic aperture radar, electro-optical/infrared, etc.) can be handled.

Finally, we expect that studies on data characteristics and adaptations will produce stronger guidance to developing more desirable data sets for evaluating research in a wider spectrum of computer vision problems.

> **DOMAIN ADAPTATION PROMISES TO BE AN ACTIVE AREA OF RESEARCH, ESPECIALLY AS ONE OF THE POSSIBLE WAYS TO QUICKLY PROPAGATE SEMANTIC ANNOTATIONS TO THE LARGE-SCALE VISUAL DATA BEING ACQUIRED EVERY MINUTE.**

AUTHORS

Vishal M. Patel (pvishalm@umd.edu) received his B.S. degrees in electrical engineering and applied mathematics (with honors) and his M.S. degree in applied mathematics from North Carolina State University, Raleigh, in 2004 and 2005, respectively. He received his Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 2010. He is a member of the research faculty at the University of Maryland Institute for Advanced Computer Studies. His research interests are in signal processing, computer vision, and machine learning with applications to imaging and biometrics. He was a recipient of the Oak Ridge Associated Universities postdoctoral fellowship in 2010. He is also a member of Eta Kappa Nu, Pi Mu Epsilon, and Phi Beta Kappa. He is a Member of the IEEE.

Raghuraman Gopalan (raghuram@research.att.com) received his Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park, in 2011. He is a senior member of technical staff at AT&T Labs–Research. His research interests are in computer vision and machine learning with a focus on object recognition problems. He is a Member of the IEEE.

Ruonan Li (ruonanli@seas.harvard.edu) received his B.E. and M.E. degrees from Tsinghua University, Beijing, China. He received the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park, in 2011. He is currently a research associate at Harvard University, Cambridge, Massachusetts. His research interests include general problems in computer vision, image processing, pattern recognition, and machine learning with recent focuses on video analysis and video-based recognition, socialized visual analytics, cross-domain model adaptation, and the application of differential geometric methods to the related problems.

Rama Chellappa (rama@umiacs.umd.edu) is a Minta Martin Professor of Engineering and the chair of the Electronics and Communication Engineering Department at the University of Maryland (UMD). He is a recipient of the K.S. Fu Prize from the International Association of Pattern Recognition, the Society, Technical Achievement and Meritorious Service Awards from the

IEEE Signal Processing Society, and the Technical Achievement and Meritorious Service Awards from the IEEE Computer Society. At UMD, he received college- and university-level recognition for research, teaching, innovation, and mentoring of undergraduate students. He is a fellow of the International Association of Pattern Recognition, the Optical Society of America, the American Association for the Advancement of Science, and the Association for Computing Machinery and holds four patents. He is a Fellow of the IEEE.

REFERENCES

[1] H. Shimodaira, "Improving predictive inference under covariate shift by weighting the log-likelihood function," *J. Stat. Plan. Inference*, vol. 90, no. 2, pp. 227–244, 2000.

[2] A. Torralba and A. A. Efros, "Unbiased look at data set bias," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2011, pp. 1521–1528.

[3] A. Khosla, T. Zhou, T. Malisiewicz, A. A. Efros, and A. Torralba, "Undoing the damage of data set bias," in *Proc. European Conf. Computer Vision*, 2012, pp. 158–171.

[4] N. Japkowicz and S. Stephen, "The class imbalance problem: A systematic study," *Intell. Data Anal.*, vol. 6, no. 5, pp. 429–450, 2002.

[5] J. J. Heckman, "Sample selection bias as a specification error," *Econometric*, vol. 47, no. 1, pp. 153–161, 1979.

[6] B. Zadrozny, "Learning and evaluating classifiers under sample selection bias," in *Proc. Int. Conf. Machine Learning*, 2004, pp. 114–121.

[7] R. Caruana, "Multitask learning," *Machine Learning*, vol. 28, no. 1, pp. 41–75, 1997.

[8] R. Raina, A. Battle, H. Lee, B. Packer, and A. Y. Ng, "Self-taught learning: Transfer learning from unlabeled data," in *Proc. Int. Conf. Machine Learning*, 2007, pp. 759–766.

[9] O. Chapelle, B. Schölkopf, and A. Zien, *Semi-Supervised Learning*. Cambridge, MA: MIT Press, 2006.

[10] A. Sharma, A. Kumar, H. Daume, and D. Jacobs, "Generalized multiview analysis: A discriminative latent space," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2012, pp. 2160–2167.

[11] J. Jiang, "Domain adaptation in natural language processing," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 2008.

[12] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY: Springer-Verlag, 1995.

[13] M. Sugiyama and K. R. Muller, "Input-dependent estimation of generalization error under covariate shift," *Stat. Decisions*, vol. 23, no. 4, pp. 249–279, 2005.

[14] J. Huang, A. J. Smola, A. Gretton, K. M. Borgwardt, and B. Schlkopf, "Correcting sample selection bias by unlabeled data," in *Advances in Neural Information Processing Systems 19*, B. Schölkopf, J. Platt, and T. Hoffman, Eds. Cambridge, MA: MIT Press, 2007, pp. 601–608.

[15] Y. Lin, Y. Lee, and G. Wahba, "Support vector machines for classification in nonstandard situations," *Mach. Learn.*, vol. 46, nos. 1–3, pp. 191–202, 2002.

[16] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowledge Data Eng.*, vol. 22, no. 10, pp. 1345–1359, 2010.

[17] W. Dai, G. Rong Xue, Q. Yang, and Y. Yu, "Transferring naive Bayes classifiers for text classification," in *Proc. AAAI Conf. Artificial Intelligence*, 2007, pp. 540–545.

[18] K. Nigam, A. K. McCallum, S. Thrun, and T. Mitchell, "Text classification from labeled and unlabeled documents using EM," *Mach. Learn.*, vol. 39, nos. 2–3, pp. 103–134, May 2000.

[19] D. Xing, W. Dai, G.-R. Xue, and Y. Yu, "Bridged refinement for transfer learning," in *Proc. European Conf. Principles and Practice of Knowledge Discovery in Databases*, 2007, pp. 324–335.

[20] J. Jiang and C. Zhai, "Instance weighting for domain adaptation in NLP," in *Proc. Annu. Meeting of the Association for Computational Linguistics*, 2007, pp. 264–271.

[21] R. Raina, "Self-taught learning," Ph.D. dissertation, Stanford University, 2009.

[22] H. Wang, F. Nie, and H. Huang, "Robust and discriminative self-taught learning," in *Proc. Int. Conf. Machine Learning*, vol. 28, no. 3, 2013, pp. 298–306.

[23] W. Dai, Q. Yang, G.-R. Xue, and Y. Yu, "Self-taught clustering," in *Proc. Int. Conf. Machine Learning, ACM*, 2008, pp. 200–207.

[24] Y. Jia, M. Salzmann, and T. Darrell, "Factorized latent spaces with structured sparsity," in *Advances in Neural Information Processing Systems*, J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, Eds. Curran Associates, Inc., 2010, pp. 982–990.

[25] N. Rasiwasia, J. Costa Pereira, E. Coviello, G. Doyle, G. R. Lanckriet, R. Levy, and N. Vasconcelos, "A new approach to cross-modal multimedia retrieval," in *Proc. Int. Conf. Multimedia*, ACM, 2010, pp. 251–260.

[26] K. Wang, R. He, W. Wang, L. Wang, and T. Tan, "Learning coupled feature spaces for cross-modal matching," in *Proc. Int. Conf. Computer Vision*, 2013, pp. 2088–2095.

[27] A. Sharma and D. Jacobs, "Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2011, pp. 593–600.

[28] Y. Jia, M. Salzmann, and T. Darrell, "Learning cross-modality similarity for multinomial data," in *Proc. IEEE Int. Conf. Computer Vision*, 2011, pp. 2407–2414.

[29] J. Yang, R. Yan, and A. G. Hauptmann, "Cross-domain video concept detection using adaptive SVMs." in *Proc. ACM Multimedia*, 2007, pp. 188–197.

[30] L. Duan, I. W. Tsang, D. Xu, and T.-S. Chua, "Domain adaptation from multiple sources via auxiliary classifiers," in *Proc. Int. Conf. Machine Learning*, 2009, pp. 289–296.

[31] H. Daumé III, "Frustratingly easy domain adaptation," in *Proc. Conf. Association for Computational Linguistics*, 2007, p. 256.

[32] W. Li, L. Duan, D. Xu, and I. Tsang, "Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 6, pp. 1134–1148, June 2014.

[33] R. Gopalan, R. Li, and R. Chellappa, "Domain adaptation for object recognition: An unsupervised approach," in *Proc. IEEE Int. Conf. Computer Vision*, 2011, pp. 999–1006.

[34] R. Gopalan, R. Li, and R. Chellappa, "Unsupervised adaptation across domain shifts by generating intermediate data representations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 11, pp. 2288–2302, Nov. 2014.

[35] J. Hoffman, B. Kulis, T. Darrell, and K. Saenko, "Discovering latent domains for multisource domain adaptation," in *Proc. European Conf. Computer Vision*, 2012, pp. 702–715.

[36] B. Gong, K. Grauman, and F. Sha, "Reshaping visual data sets for domain adaptation," in *Proc. Neural Information Processing Systems*, 2013, pp. 1286–1294.

[37] B. Gong, Y. Shi, F. Sha, and K. Grauman, "Geodesic flow kernel for unsupervised domain adaptation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2012, pp. 2066–2073.

[38] B. Gong, K. Grauman, and F. Sha, "Learning kernels for unsupervised domain adaptation with applications to visual object recognition," *Int. J. Comput. Vis.*, vol. 109, no. 1–2, pp. 3–27, Aug. 2014.

[39] J. Zheng, M.-Y. Liu, R. Chellappa, and P. Phillips, "A Grassmann manifold-based domain adaptation approach," in *Proc. Int. Conf. Pattern Recognition*, 2012, pp. 2095–2099.

[40] A. Shrivastava, S. Shekhar, and V. M. Patel, "Unsupervised domain adaptation using parallel transport on Grasman manifold," in *Proc. IEEE Winter Conf. Applications of Computer Vision*, 2014, pp. 277–284.

[41] K. Saenko, B. Kulis, M. Fritz, and T. Darrell, "Adapting visual category models to new domains," in *Proc. European Conf. Computer Vision*, 2010, vol. 6314, pp. 213–226.

[42] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon, "Information-theoretic metric learning," in *Proc. Int. Conf. Machine Learning*, 2007, pp. 209–216.

[43] B. Kulis, K. Saenko, and T. Darrell, "What you saw is not what you get: Domain adaptation using asymmetric kernel transforms," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2011, pp. 1785–1792.

[44] I.-H. Jhuo, D. Liu, D. Lee, and S.-F. Chang, "Robust visual domain adaptation with low-rank reconstruction," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2012, pp. 2168–2175.

[45] M. Baktashmotlagh, M. T. Harandi, B. C. Lovell, and M. Salzmann, "Unsupervised domain adaptation by domain invariant projection," in *Proc. IEEE Int. Conf. Computer Vision*, 2013, pp. 769–776.

[46] B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars, "Unsupervised visual domain adaptation using subspace alignment," in *Proc. IEEE Int. Conf. Computer Vision*, 2013, pp. 2960–2967.

[47] Y. Aytar and A. Zisserman, "Tabula rasa: Model transfer for object category detection," in *Proc. IEEE Int. Conf. Computer Vision*, 2011, pp. 2252–2259.

[48] W. Jiang, E. Zavesky, S.-F. Chang, and A. Loui, "Cross-domain learning methods for high-level visual concept classification," in *Proc. IEEE Int. Conf. Image Processing*, 2008, pp. 161–164.

[49] L. Duan, I. W.-H. Tsang, D. Xu, and S. J. Maybank, "Domain transfer SVM for video concept detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009, pp. 1375–1381.

[50] A. Bergamo and L. Torresani, "Exploiting weakly-labeled web images to improve object classification: A domain adaptation approach," in *Advances in Neural Information Processing Systems*, J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, Eds. Curran Associates, Inc., 2010, pp. 181–189.

[51] L. Bruzzone and M. Marconcini, "Domain adaptation problems: A DASVM classification technique and a circular validation strategy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 770–787, 2010.

[52] L. Duan, D. Xu, and I. Tsang, "Domain adaptation from multiple sources: A domain-dependent regularization approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 3, pp. 504–518, 2012.

[53] A. J. Ma, P. C. Yuen, and J. Li, "Domain transfer support vector ranking for person re-identification without target camera label information," in *Proc. IEEE Int. Conf. Computer Vision*, 2013, pp. 3567–3574.

[54] L. Duan, D. Xu, and S.-F. Chang, "Exploiting web images for event recognition in consumer videos: A multiple source domain adaptation approach," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2012, pp. 1338–1345.

[55] M. Gönen and E. Alpaydin, "Multiple kernel learning algorithms," *J. Mach. Learn. Res.*, vol. 12, pp. 2211–2268, July 2011.

[56] L. Duan, D. Xu, I.-H. Tsang, and J. Luo, "Visual event recognition in videos by learning from web data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1667–1680, 2012.

[57] L. Duan, I. Tsang, and D. Xu, "Domain transfer multiple kernel learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 3, pp. 465–479, 2012.

[58] Z. Guo and Z. J. Wang, "Cross-domain object recognition via input-output kernel analysis," *IEEE Trans. Image Processing*, vol. 22, no. 8, pp. 3108–3119, Aug. 2013.

[59] Y. Shi and F. Sha, "Information-theoretical learning of discriminative clusters for unsupervised domain adaptation," in *Proc. Int. Conf. Machine Learning*, 2012, pp. 1079–1086.

[60] J. Hoffman, E. Rodner, J. Donahue, T. Darrell, and K. Saenko, "Efficient learning of domain-invariant image representations," in *Proc. Int. Conf. Learning Representations*, 2013, pp. 1–9.

[61] J. Donahue, J. Hoffman, E. Rodner, K. Saenko, and T. Darrell, "Semi-supervised domain adaptation with instance constraints," in *Proc. IEEE Int. Conf. Computer Vision*, 2013, pp. 668–675.

[62] B. A. Olshausen and D. J. Fieldt, "Sparse coding with an overcomplete basis set: A strategy employed by v1," *Vis. Res.*, vol. 37, no. 23, pp. 3311–3325, Dec. 1997.

[63] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. S. Huang, and S. Yan, "Sparse representation for computer vision and pattern recognition," *Proc. IEEE*, vol. 98, no. 6, pp. 1031–1044, 2010.

[64] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. IEEE*, vol. 98, no. 6, pp. 1045–1057, June 2010.

[65] Q. Qiu, V. M. Patel, P. Turaga, and R. Chellappa, "Domain adaptive dictionary learning," in *Proc. European Conf. Computer Vision*, 2012, vol. 7575, pp. 631–645.

[66] S. Shekhar, V. M. Patel, H. V. Nguyen, and R. Chellappa, "Generalized domain-adaptive dictionaries," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2013, pp. 361–368.

[67] J. Ni, Q. Qiu, and R. Chellappa, "Subspace interpolation via dictionary learning for unsupervised domain adaptation," in *Proc. IEEE Int. Conf. Computer Vision*, 2013, pp. 692–699.

[68] J. Zheng, R. Chellappa, and P. J. Phillips, "Sparse embedding-based domain adaptation for object recognition," in *Proc. IEEE Int. Conf. Computer Vision Workshop on Visual Domain Adaptation and Dataset Bias*, 2013, pp. 1–2.

[69] H. V. Nguyen, V. M. Patel, N. M. Nasrabadi, and R. Chellappa, "Sparse embedding: A framework for sparsity promoting dimensionality reduction," in *Proc. European Conf. Computer Vision*, 2012, pp. 414–427.

[70] B. Gong, K. Grauman, and F. Sha, "Connecting the dots with landmarks: Discriminatively learning domain-invariant features for unsupervised domain adaptation," in *Proc. Int. Conf. Machine Learning*, 2013, pp. 222–230.

[71] B. Gong, F. Sha, and K. Grauman, "Overcoming data set bias: An unsupervised domain adaptation approach," in *Proc. Neural Information Processing Systems Workshop on Large Scale Visual Recognition and Retrieval*, 2012, pp. 1–5.

[72] Y. Bengio, "Learning deep architectures for AI," *Found. Trends Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.

[73] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, July 2006.

[74] G. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[75] M. Chen, Z. Xu, and K. Q. Weinberger, "Marginalized denoising autoencoders for domain adaptation," in *Proc. Int. Conf. Machine Learning*, 2012, pp. 767–774.

[76] X. Glorot, A. Bordes, and Y. Bengio, "Domain adaptation for large-scale sentiment classification: A deep learning approach," in *Proc. Int. Conf. Machine Learning*, 2011, pp. 513–520.

[77] H. Van Nguyen, "Non-linear and sparse representations for multi-modal recognition," Ph.D. dissertation, Univ. Maryland, 2013.

[78] S. Chopra, S. Balakrishnan, and R. Gopalan, "DLID: Deep learning for domain adaptation by interpolating between domains," in *Proc. ICML Workshop on Challenges in Representation Learning*, 2013, pp. 1–8.

[79] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition," in *Proc. Int. Conf. Machine Learning*, 2013, pp. 647–655.

[80] L. Bo, X. Ren, and D. Fox, "Hierarchical matching pursuit for image classification: Architecture and fast algorithms," in *Proc. Neural Information Processing Systems*, 2011, pp. 2115–2123.

[81] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and simile classifiers for face verification," in *Proc. IEEE Int. Conf. Computer Vision*, Oct. 2009, pp. 365–372.

[82] A. Farhadi, I. Endres, D. Hoiem, and D. Forsyth, "Describing objects by their attributes," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009, pp. 1778–1785.

[83] Y. Wang and G. Mori, "A discriminative latent model of object classes and attributes," in *Proc. European Conf. Computer Vision*, 2010, pp. 155–168.

[84] S. Branson, C. Wah, B. Babenko, F. Schroff, P. Welinder, P. Perona, and S. Belongie, "Visual recognition with humans in the loop," in *Proc. European Conf. Computer Vision*, pp. 438–451.

[85] C. H. Lampert, H. Nickisch, and S. Harmeling, "Learning to detect unseen object classes by between-class attribute transfer," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009, pp. 951–958.

[86] A. Kovashka and K. Grauman, "Attribute adaptation for personalized image search," in *Proc. IEEE Int. Conf. Computer Vision*, 2013, pp. 3432–3439.

[87] T. Tommasi and B. Caputo, "Frustratingly easy NBNN domain adaptation," in *Proc. IEEE Int. Conf. Computer Vision*, 2013, pp. 897–904.

[88] S. S. F. V. Jain, "Adapting classification cascades to new domains," in *Proc. IEEE Int. Conf. Computer Vision*, 2013, pp. 105–112.

[89] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," in *ACM Computing Surveys*, Dec. 2003, pp. 399–458.

[90] H. Ho and R. Gopalan, "Model-driven domain adaptation on product manifolds for unconstrained face recognition," *Int. J. Comput. Vis.*, vol. 109, no. 1–2, pp. 110–125, Aug. 2014.

[91] R. Gross, I. Matthews, J. F. Cohn, T. Kanade, and S. Baker, "Multi-pie," *Image Vis. Comput.*, vol. 28, no. 5, pp. 807–813, 2010.

[92] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Fisher discrimination dictionary learning for sparse representation," in *Proc. Int. Conf. Computer Vision*, 2011, pp. 543–550.

[93] G. Griffin, A. Holub, and P. Perona. (2007). Caltech-256 object category data set. California Institute of Technology, Tech. Rep. 7694. [Online]. Available: http://authors.library.caltech.edu/7694

[94] K. Gallivan, A. Srivastava, X. Liu, and P. Van Dooren, "Efficient algorithms for inferences on Grassmann manifolds," in *Proc. IEEE Workshop Statistical Signal Processing*, Sept. 2003, pp. 315–318.

[95] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the k-svd algorithm using batch orthogonal matching pursuit," Technion–Computer Science Department, Tech. Rep. CS-2008-08, Apr. 2008.

[96] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York: Springer, 2010.

[97] A. Yang, Z. Zhou, A. Balasubramanian, S. Sastry, and Y. Ma, "Fast l1-minimization algorithms for robust face recognition," *IEEE Trans. Image Processing*, vol. 22, no. 8, pp. 3234–3246, Aug. 2013.

[98] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.

[99] Z. Lin, A. Ganesh, J. Wright, L. Wu, M. Chen, and Y. Ma, "Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix," UIUC Technical Report, Tech. Rep. UILU-ENG-09-2214, July 2009.

[100] Z. Lin, M. Chen, L. Wu, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," UIUC Tech. Rep. UILU-ENG-09-2215, Oct 2009.

[101] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge Univ. Press, 2004.

[102] X. Wang and X. Tang, "Face photo-sketch synthesis and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 1955–1967, Nov. 2009.

[103] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: a database for studying face recognition in unconstrained environments," Univ. Massachusetts, Amherst, LFWTech. Rep. 07-49, Oct. 2007.

[SP]

IEEE SIGNAL PROCESSING MAGAZINE [69] MAY 2015

SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page
Qmags THE WORLD'S NEWSSTAND®

[Tiancheng Li, Miodrag Bolić, and Petar M. Djurić]

Resampling Methods for Particle Filtering

[Classification, implementation, and strategies]

IMAGE LICENSED BY GRAPHIC STOCK

Two decades ago, with the publication of [1], we witnessed the rebirth of particle filtering (PF) as a methodology for sequential signal processing. Since then, PF has become very popular because of its ability to process observations represented by nonlinear state-space models where the noises of the model can be non-Gaussian. This methodology has been adopted in various fields, including finance, geophysical systems, wireless communications, control, navigation and tracking, and robotics [2]. The popularity of PF has also spurred the publication of several review articles [2]–[6].

Using the PF method, we aim to track various distributions that arise in dynamic state-space models. The tracking is carried out by exploring the space of the states with randomly generated samples (also called *particles*). The

distributions of interest are approximated by the generated particles as well as weights assigned to the particles.

There are many PF methods, and almost all of them are based on three operations: 1) particle propagation, 2) weight computation, and 3) resampling. Particle propagation and weight computation amount to the generation of particles and assignment of weights, whereas resampling replaces one set of particles and their weights with another set.

Particle generation and weight computation are computationally the most intensive steps. However, they are application dependent and can be easily implemented in parallel if parallel hardware is available. The resampling step is universal and generally state-dimension-free but is not naturally suitable for parallel processing. The resampling is essential for PF; without this step, PF will quickly produce a degenerate set of particles, i.e., a set in which a few particles dominate the rest of the particles with their weights. This means that the obtained estimates will be

inaccurate and will have unacceptably large variances. With resampling, such deteriorations are prevented, which is why it is highly important to PF. Consequently, resampling has been extensively researched, and, as a result, various resampling schemes have been proposed [7]–[10].

Surveys of resampling methods can be found in [11]–[14]. These papers, however, only cover a small number of basic resampling methods and have become somewhat outdated. Furthermore, no classification has been made in these papers. This has been a disadvantage to researchers in getting a better grasp of the overall picture of resampling and of being able to readily choose the scheme that would fit their needs. This article aims to correct these shortfalls. More specifically, the main goal of this article is to introduce a new classification of resampling algorithms and provide a qualitative comparison of them. An additional goal is to set the grounds for further developments in resampling.

We divide the resampling algorithms into sequential and distributed/parallel algorithms. The most common are the sequential algorithms, where resampling is carried out from the approximating distribution using the latest weights. We also refer to these algorithms as *traditional*. However, researchers have explored other solutions, such as resampling that takes into account the history of the weights, resampling from approximate distributions, and resampling from only a part of the sampling space. In addition, approaches based on different ways of grouping particles have been proposed where grouping can be implemented, e.g., by using thresholds or by combining adjacent particles. These approaches might be of use when it is important to reduce the number of operations or to reduce communication between processing elements (PEs) in parallel implementations. We also discuss the frequency of resampling.

With multicore processors, the existence of general-purpose graphical processing units (GP-GPUs) in almost every computer and the emergence of embedded multicore and GP-GPU hardware, parallel processing can be readily implemented. The implementations of parallel resampling, however, can vary, which, in turn, reflect on their accuracy and speed. Some of the efforts in this area are described in this article.

In our view, future research efforts will be directed to specific implementations and theoretical analysis of the methods. The former include simplifying the resampling algorithms, development of better schemes with the ultimate goal of improving filtering performance, parallelization, and real-time implementations. The latter addresses the effects of the resampling algorithms on convergence and accuracy of approximation. As this article focuses mainly on providing guidelines to the readers and on qualitative description of the solutions, we do not explore issues such as robustness and do not provide theoretical proofs of any sort.

BACKGROUND OF PF AND RESAMPLING

A BRIEF REVIEW OF PF
We start with a brief review of PF and introduce the notation. There is a state-space model described by

$$x_t = g(x_{t-1}, u_t), \qquad (1)$$
$$y_t = h(x_t, v_t), \qquad (2)$$

where t is a time index and $t = 1, 2, \ldots;$ $x_t \in R^{d_x}$ is the state of the model that is hidden (not observed); $y_t \in R^{d_y}$ is the observation; $u_t \in R^{d_u}$ and $v_t \in R^{d_v}$ are white noises that are independent of each other; and $g: R^{d_x} \times R^{d_u} \to R^{d_x}$ and $h: R^{d_x} \times R^{d_v} \to R^{d_y}$ are known functions. An alternative representation of these equations is by the probability distributions of the state, $p(x_t | x_{t-1})$, and of the observation, $p(y_t | x_t)$, which can be obtained from (1) and (2) and the probability distributions of u_t and v_t, respectively. The interest is in nonlinear models and where the noises in (1) and (2) are not necessarily Gaussian.

The objective of PF is the sequential estimation of distributions of the state, including the filtering distribution $p(x_t | y_{1:t})$, the predictive distribution $p(x_t | y_{1:t-1})$, or the smoothing distribution $p(x_t | y_{1:T})$, where $t < T$. Here, we focus on the filtering distribution. This distribution can be expressed in terms of the filtering distribution at time instant $t - 1$, $p(x_{t-1} | y_{1:t-1})$, i.e., in a recursive form by

$$p(x_t | y_{1:t}) \propto \int p(y_t | x_t)\, p(x_t | x_{t-1})\, p(x_{t-1} | y_{1:t-1}) dx_{t-1}, \qquad (3)$$

where the symbol $\propto$ signifies "proportional to." This update cannot be implemented analytically except in a very few cases, and, therefore, one resorts to approximations.

We reiterate that with PF, the underlying approximation is to represent continuous distributions by discrete random measures composed of particles $x_t^{(m)}$, which are possible values of the unknown state x_t and weights $w_t^{(m)}$ assigned to the particles. The distribution $p(x_{t-1} | y_{1:t-1})$ is approximated by a random measure of the form $\chi_{t-1} = \{x_{t-1}^{(m)}, w_{t-1}^{(m)}\}_{m=1}^M$, where M is the number of particles, i.e.,

$$p(x_{t-1} | y_{1:t-1}) \approx \sum_{m=1}^M w_{t-1}^{(m)} \delta(x_{t-1} - x_{t-1}^{(m)}), \qquad (4)$$

where $\delta(\cdot)$ is the Dirac delta impulse and all the weights sum up to one. With this approximation, the integral in (3) can readily be solved, and we can write

$$p(x_t | y_{1:t}) \dot{\propto} p(y_t | x_t) \sum_{m=1}^M w_{t-1}^{(m)} p(x_t | x_{t-1}^{(m)}), \qquad (5)$$

where $\dot{\propto}$ means "approximate proportionality."

The last expression shows how we can obtain the approximation χ_t of the filtering distribution recursively in time. At time instant $t - 1$, one starts the construction of χ_t by generating particles $x_t^{(m)}$, which are used for representing $p(x_t | y_{1:t})$. This step of PF is referred to as *particle propagation* because a particle $x_{t-1}^{(m)}$ is moved forward in time and is a parent of $x_t^{(m)}$. For particle propagation and weight computation, we employ the concept of importance sampling [15]. Ideally, the propagated particles should be drawn from $p(x_t | y_{1:t})$, and then they will all have equal weights. However, this is infeasible in most cases, and,

therefore, one uses an instrumental function $\pi(x_t)$. For example, in [1], this function is $p(x_t | x_{t-1})$.

The second basic step of PF is the computation of the particle weights. To have correct inference from the generated particles, the theory shows that the generated particles from $\pi(x_t)$, which is different from $p(x_t | y_{1:t})$, have to be weighted [15]. Under mild conditions, one can show that these weights can be recursively computed according to

$$w_t^{(m)} \propto w_{t-1}^{(m)} \frac{p(y_t | x_t^{(m)}) \, p\left(x_t^{(m)} | x_{t-1}^{(m)}\right)}{\pi(x_t^{(m)})}. \qquad (6)$$

Often, the computation of the expression to the right of the proportionality sign is followed by normalization of the weights (so that they sum up to one).

Ideally, the weights of the particles should all be equal. On the other extreme, it is most undesirable if all the particles have weights equal to zero or one or a few particles have most of the weight and the rest of the particle weights are negligible. This is commonly called *degeneracy* and is exactly what eventually happens when PF is realized by using only the aforementioned two steps. Then, as the processing of the observations proceeds, the variance of the weights increases and reaches a point at which the random measure is a very poor approximation of the filtering distribution. For this reason, PF needs a third step, referred to as *resampling*.

THE BASICS OF RESAMPLING

With resampling, one aims to prevent the degeneracy of the propagated particles by modifying the random measure $\mathcal{X}_t$ to $\tilde{\mathcal{X}}_t$ and improving the exploration of the state space at $t+1$. While alleviating degeneracy during resampling, it is important that the random measure approximates the original distribution as well as possible and prevents bias in the estimates [16]. Although the approximation by $\tilde{\mathcal{X}}_t$ is very similar to that of $\mathcal{X}_t$, the set of particles of $\tilde{\mathcal{X}}_t$ is significantly different from that of $\mathcal{X}_t$. Resampling means that the particles from $\mathcal{X}_t$ with large weights are more likely to dominate $\tilde{\mathcal{X}}_t$ than particles with small weights, and, consequently, in the next time step, more new particles will be generated in the region of large weights. This is the reason for the improvement in exploration after resampling. The focus of exploration is shifted to the parts of the space with large probability masses. Because of resampling, the propagated particles from $\tilde{\mathcal{X}}_t$ will have weights that are less discriminate than if the propagation was from the particles of $\mathcal{X}_t$. This is an intuitive idea with important practical and theoretical implications.

Formally, resampling is a process in which one samples from the original random measure $\mathcal{X}_t = \{x_t^{(m)}, w_t^{(m)}\}_{m=1}^{M}$ to create a new random measure $\tilde{\mathcal{X}}_t = \{\tilde{x}_t^{(n)}, \tilde{w}_t^{(n)}\}_{n=1}^{N}$. Then, the random measure $\mathcal{X}_t$ is replaced with $\tilde{\mathcal{X}}_t$. In the process, some of the particles of $\mathcal{X}_t$ are replicated, and, most likely, they are the particles with large weights. The particles of $\tilde{\mathcal{X}}_t$ are used for propagation of new particles, and thus, they are the parents of $x_{t+1}^{(m)}$. We note that for the approximation of $p(x_t | y_{1:t})$, it is better to use $\mathcal{X}_t$ than $\tilde{\mathcal{X}}_t$. We also observe that the number of resampled particles N is not necessarily equal to the number of propagated particles. Traditional resampling methods keep them constant, and, usually, $M = N$. Finally, in most of the resampling methods, the weights of the particles after resampling are all equal.

Resampling, however, may introduce undesired effects. One of them is sample impoverishment. With resampling, low-weighted particles are most likely removed, and, thereby, the diversity of the particles is reduced [1], [9], [12]. For example, if a few particles of $\mathcal{X}_t$ have most of the weight, many of the resampled particles will be the same, i.e., the number of different particles in $\tilde{\mathcal{X}}_t$ will be small. The other effect is on the speedup of implementation of PF. We recall that PF is often used for signal processing when observations have to be processed in real time. A good solution for gaining in speed is to parallelize the PF. As will be shown, parallelizing the resampling is rather challenging.

> WHEN IMPLEMENTING RESAMPLING, SEVERAL DECISIONS MUST BE MADE. THEY INCLUDE CHOOSING THE DISTRIBUTION FOR RESAMPLING, SPECIFYING THE SAMPLING STRATEGY, DETERMINING THE RESAMPLED SIZE, AND SELECTING THE FREQUENCY OF RESAMPLING.

The undesired effects of resampling have pushed researchers to investigate advanced methods for resampling. These methods have many features, including a variable number of particles, the removal of the restriction of equal weighting of the resampled particles, the avoidance of discarding low-weighted particles, and the introduction of parallel frameworks for resampling.

When implementing resampling, several decisions must be made. They include choosing the distribution for resampling, specifying the sampling strategy, determining the resampled size, and selecting the frequency of resampling.

CLASSIFICATION OF RESAMPLING SCHEMES

Our high-level classification of resampling methods with representative references is shown in Table 1. We first group the methods based on their implementation as sequential and parallel. We note that the parallel implementations represent two or more sequential implementations executed simultaneously. The sequential strategies are further classified based on whether the resampling is from a single distribution or from two or more distributions obtained from grouping of the particles (compound sampling). We also have a third category, referred to as *special strategies*. As the name suggests, the special strategies have features that separate them from single and compound sampling. Next, we make several remarks, of which the first four are also reflected in Table 1 (R1–R4).

[TABLE 1] THE HIGH-LEVEL CLASSIFICATION OF RESAMPLING METHODS.

CLASSIFICATIONS	R1 BASED ON THE DISTRIBUTION $\mathcal{X}_t$	R2 RESAMPLING OF ALL THE PARTICLES IN THE SAME WAY	R3 GROUPING	R4 USING INFORMATION FROM THE CURRENT TIME INSTANT
Sequential implementation	YES	YES/NO	YES/NO	YES/NO
■ SINGLE DISTRIBUTION SAMPLING [1], [7], [8], [10]	YES	YES	NO	YES
■ COMPOUND-SAMPLING		YES/NO FOR MANY: DIFFERENT OR NO RESAMPLING PER GROUP	BASED ON COMPOUND CRITERIA	
• THRESHOLDS/GROUPING-BASED RESAMPLING [11], [12], [24]				YES/NO FOR [11] AND [17]
• RESAMPLING THAT TAKES INTO ACCOUNT THE STATE [16]				YES
■ SPECIAL STRATEGIES		YES		
• MODIFIED RESAMPLING [11]	NO		NO	YES/NO FOR [18]
• VARIABLE-SIZE RESAMPLING [27]	YES			YES
• ROUGHENING [1]				
Parallel implementation			BASED ON ADJACENT PARTICLES	
■ MAPPING TO SPECIFIC HARDWARE PLATFORMS [31]–[33], [35]–[38]				
■ DISTRIBUTED RESAMPLING [32], [40]				
■ NORMALIZATION-FREE RESAMPLING [38], [39], [41]				

REMARK 1

One classification is based on the used distribution for resampling. Mainly, this is the distribution represented by $\mathcal{X}_t$. Other approaches include sampling from an approximate distribution. We will point out cases in which resampling is not performed from $\mathcal{X}_t$.

REMARK 2

Resampling can be performed in the same way on all of the particles, or it is possible to resample in different ways in different parts of the sampling space.

REMARK 3

There are methods that group the particles in some way before resampling is performed. In that respect, we distinguish among single-distribution sampling schemes (no grouping of particles), techniques that combine adjacent particles (which is the common approach in parallel implementations of resampling), and techniques that group particles for resampling based on some predefined criteria (referred to as *compound sampling*).

REMARK 4

Resampling can also be classified based on whether only particles from the current time step are involved in resampling, which is a common approach, the particles from previous time instants are considered [17], or some future particles are generated and taken into account for resampling [18]. Also, resampling can be classified based on whether only the weight is taken into account, which is the standard approach, or the state and the weight of particles are considered together [16].

REMARK 5

Resampling may be applied not in all but only in selected time steps. Compensations such as roughening [1] may be implemented with resampling for performance improvement.

REMARK 6

There are deterministic and stochastic resampling methods. The deterministic methods lead to a set of particles that are always the same for the same input set of particles.

IMPLEMENTATION OF RESAMPLING SCHEMES

In the following sections, we will explain various resampling methods and provide pseudocodes for selected algorithms. The pseudocodes are presented in a simple and unified way but not in forms that optimize the implementation of the algorithms. In practical implementations on a specific platform, programming techniques are required to maximally speed up the computation. For example, the cumulative sum calculation in code 1 can be implemented in MATLAB using vectors, which would be faster than the iterative calculation shown by code 1. Traditional sampling methods have already been described elsewhere; we present them here for completeness and because they are used in the compound, special, and parallel methods described later. We assume that the weights $w_t^{(m)}$ are normalized before resampling, i.e., $\sum_{m=1}^{M} w_t^{(m)} = 1$.

SINGLE-DISTRIBUTION SAMPLING METHODS

In this category, all the particles are resampled by using a single-distribution sampling procedure. The expected number of times $N_t^{(m)}$ that the mth particle is resampled is proportional to $w_t^{(m)}$, i.e.,

[TABLE 2] PSEUDOCODES OF MULTINOMIAL, STRATIFIED, SYSTEMATIC, RESIDUAL, BRANCH-KILL, AND ROUNDING-COPY RESAMPLING.

Code 1: Cumulative sum of normalized weights.

$[\{Q_t^{(m)}\}_{m=1}^M] = \text{CumulativeSum}[\{w_t^{(m)}\}_{m=1}^M]$

$Q_t^{(1)} = w_t^{(1)}$
FOR $m = 2:M$
    $Q_t^{(m)} = Q_t^{(m-1)} + w_t^{(m)}$
END

Code 2: Deterministic replication of particles.

$[\{\bar{x}_t^{(n)}\}_{n=1}^N, N] = \text{Replication}[\{x_t^{(m)}, N_t^{(m)}\}_{m=1}^M]$

$n = 0$
FOR $m = 1:M$
    FOR $h = 1:N_t^{(m)}$
        $n = n + 1$
        $\bar{x}_t^{(n)} = x_t^{(m)}$
    END
END
$N = n$

Code 3: Multinomial/stratified/systematic resampling.

$[\{\bar{x}_t^{(n)}\}_{n=1}^N] = \text{Resample}[\{x_t^{(m)}, w_t^{(m)}\}_{m=1}^M, N]$

$[\{Q_t^{(m)}\}_{m=1}^M] = \text{CumulativeSum}[\{w_t^{(m)}\}_{m=1}^M]$
$n = 0$
/**Systematic**/**stratified** choice runs:
    $m = 1$
/**Systematic** choice runs
    $u_0 \sim U(0, 1/N]$
WHILE $(n \leq N)$
/**Stratified** choice runs
    $u_0 \sim U(0, 1/N]$
/**Systematic**/**stratified** choice runs
    $u = u_0 + n/N$
/**Multinomial** choice runs
    $u \sim U(0, 1]; \quad m = 1$
    WHILE $(Q_t^{(m)} < u)$
        $m = m + 1$
    END
    $n = n + 1$
    $\bar{x}_t^{(n)} = x_t^{(m)}$
END

Code 4: Residual resampling/residual systematic resampling (RSR).

$[\{\bar{x}_t^{(n)}\}_{n=1}^N] = \text{Resample}[\{x_t^{(m)}, w_t^{(m)}\}_{m=1}^M, N]$

/**RSR** choice runs
    $\Delta u \sim U(0, 1/N]$
FOR $m = 1:M$
/**Residual** choice runs the following two lines
    $N_t^{(m)} = \text{Floor}(N \times w_t^{(m)})$
    $\hat{w}_t^{(m)} = w_t^{(m)} - N_t^{(m)}/N$
/**RSR** choice runs the following two lines
    $N_t^{(m)} = \text{Floor}(N \times (w_t^{(m)} - \Delta u)) + 1$
    $\Delta u = \Delta u + N_t^{(m)}/N - w_t^{(m)}$
END
$[\{\bar{x}_t^{(n)}\}_{n=1}^{N_t}, N_t] = \text{Replication}[\{x_t^{(m)}, N_t^{(m)}\}_{m=1}^M]$
/**Residual** choice runs the following four lines
    FOR $m = 1:M$
        $\hat{w}_t^{(m)} = \hat{w}_t^{(m)} \times N/(N - N_t)$
    END
    $[\{\bar{x}_t^{(n)}\}_{n=N_t+1}^N] = \text{(Multinomial)Resample}[\{x_t^{(m)}, \hat{w}_t^{(m)}\}_{m=1}^M, N - N_t]$

Code 5: Branch-kill/rounding-copy resampling.

$[\{\bar{x}_t^{(n)}\}_{n=1}^N] = \text{Resample}[\{x_t^{(m)}, w_t^{(m)}\}_{m=1}^M, N_{\text{ref}}]$

FOR $m = 1:M$
/**Branch kill** choice runs the following five lines
    $u \sim U(0, 1/N_{\text{ref}}]$
    $N_t^{(m)} = \text{Floor}(N_{\text{ref}} \times w_t^{(m)})$
    IF $(N_{\text{ref}} \times w_t^{(m)} - N_t^{(m)}) \geq u$
        $N_t^{(m)} = N_t^{(m)} + 1$
    END
/**Rounding-copy** choice runs
    $N_t^{(m)} = \text{Round}(N_{\text{ref}} \times w_t^{(m)})$
END
$[\{\bar{x}_t^{(n)}\}_{n=1}^N, N] = \text{Replication}[\{x_t^{(m)}, N_t^{(m)}\}_{m=1}^M]$

Note: The different methods are described with different colors, and the black text is common for all of them in each group, except where otherwise stated. Code 3 presents resampling methods based on the cumulative sum of the normalized weights (obtained by code 1). Codes 4 and 5 are methods that use deterministic replication (implemented by code 2). All of these codes produce resampled particles with equal weights.

$$E(N_t^{(m)} | w_t^{(m)}) = Nw_t^{(m)}. \tag{7}$$

This constraint is known as the *unbiasedness* or *proper-weighting condition* [10].

TRADITIONAL METHODS

Multinomial Resampling

The core idea of multinomial resampling [1] is to generate independently N random numbers, $u_t^{(n)}$ from the uniform distribution on (0, 1] and use them to select particles from χ_t. In the nth selection, the particle $x_t^{(m)}$ is chosen when the following condition is satisfied:

$$Q_t^{(m-1)} < u_t^{(n)} \leq Q_t^{(m)}, \tag{8}$$

where

$$Q_t^{(m)} = \sum_{k=1}^m w_t^{(k)}. \tag{9}$$

Thus, the probability of selecting $x_t^{(m)}$ is the same as that of $u_t^{(n)}$ being in the interval bounded by the cumulative sum of the normalized weights as shown in (8). This sampling scheme satisfies the unbiasedness condition. The pseudocode of the cumulative sum of normalized weights is shown by code 1 in Table 2.

Multinomial resampling (see code 3 in Table 2) is also referred to as *simple random resampling*. Since the sampling of each particle is random, the upper and lower limits of the number of times a given particle is resampled are zero (not sampled) and N_t (sampled N_t times), respectively. This yields the maximum variance of the resampled particles.
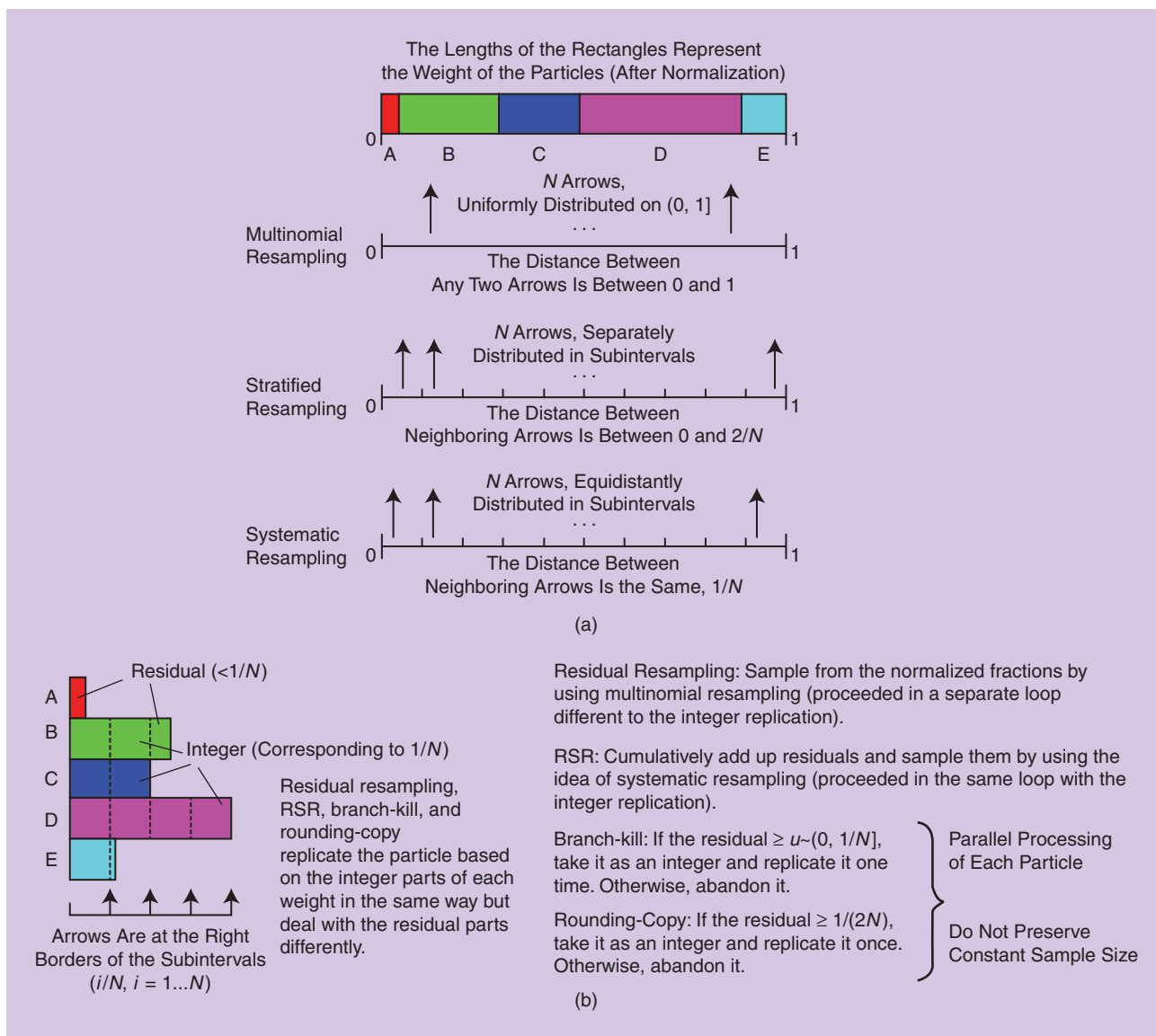
[FIG1] A description of the traditional resampling methods. (a) A multinomial, stratified, and systematic resampling based on the cumulative sum of the normalized particle weights. The arrows represent sampling locations, and a particle is sampled if it is targeted by an arrow. (b) Methods based on residual resampling. They include two parts. The particles from the first part are obtained by replicating $x_t^{(m)}$ $\lfloor Nw_t^{(m)} \rfloor$ times if $\lfloor Nw_t^{(m)} \rfloor \geq 1$, and the particles from the second part are generated based on the residual weights.

The computational complexity of multinomial resampling is of order $O(NM)$, where the M factor arises from the search of the required m in (8). It is known that multinomial resampling is not efficient [8], and this has motivated a search for faster methods. The binary search is explored to execute the search of m in (8), which reduces the computational complexity from M to $\log(M)$ [8]. The variance of the number of times a particle is resampled can be reduced by, for example, stratification and deterministic sampling.

Stratified/Systematic Resampling

Stratified resampling [7] divides the whole population of particles into subpopulations called *strata*. It prepartitions the $(0, 1]$ interval into N disjoint subintervals $(0, 1/N] \cup \cdots \cup (1 - 1/N, 1]$. The random

numbers $\{u_t^{(n)}\}_{n=1}^N$ are drawn independently in each of these subintervals, i.e.,

$$u_t^{(n)} \sim U\left(\frac{n-1}{N}, \frac{n}{N}\right], \qquad n = 1, 2, \ldots, N, \qquad (10)$$

and then the bounding method based on the cumulative sum of normalized weights as shown in (8) is used.

Systematic resampling [7], [8] also exploits the idea of strata but in a different way. Now, $u_t^{(1)}$ is drawn from the uniform distribution on $(0, 1/N]$, and the rest of the u numbers are obtained deterministically, i.e.,

$$u_t^{(1)} \sim U\left(0, \frac{1}{N}\right],$$

$$u_t^{(n)} = u_t^{(1)} + \frac{n-1}{N}, \qquad n = 2, 3, \ldots, N. \qquad (11)$$

Both the stratified and the systematic procedures are presented in Table 2. Their complexity is of order $O(N)$. Note that the systematic method is computationally more efficient than the stratified method because of the smaller number of random numbers that are generated. A visual description of the multinomial, stratified, and systematic resampling methods is displayed in Figure 1(a).

The upper and lower limits of the times the mth particle is resampled in the systematic method are $\lfloor N w_t^{(m)} \rfloor$ and $\lfloor N w_t^{(m)} \rfloor + 1$, respectively, where $\lfloor x \rfloor$ denotes the floor operation (the largest integer not exceeding x). By contrast, for stratified resampling, they are $\max(\lfloor N w_t^{(m)} \rfloor - 1, 0)$ and $\lfloor N w_t^{(m)} \rfloor + 2$, respectively, because the variables $\{u_t^{(n)}\}_{n=1}^N$ are not equidistant, and instead $\Delta u = u_t^{(n)} - u_t^{(n-1)}$ for $n = 2, 3, \ldots, N$ varies between 0 and $2/N$. When Δu is close to 0, a particle with a very small weight (close to 0 but bigger than Δu) can be resampled twice and when Δu is bigger than $2/N$, a particle with weight between $1/N$ and Δu can be discarded. This indicates that the variance of the number of a resampled particle by the systematic method is smaller than that of the stratified method.

Residual Resampling (Remainder Resampling)
Residual resampling [9], [10] consists of two stages. The first is a deterministic replication of each particle whose weight is bigger than $1/N$, and the second is random sampling using the remaining of the weights (referred to as *residuals*). The code for deterministic replication is shown by code 2, where $N_t^{(m)}$ represents the number of times the particle $x_t^{(m)}$ is replicated in this way. With residual resampling, the mth particle is resampled $N_t^{(m)} + R_t^{(m)}$ times, where $N_t^{(m)}$ and $R_t^{(m)}$ are the numbers of replications from the first and second stage, respectively, and where $N_t^{(m)} = \lfloor N w_t^{(m)} \rfloor$. The total number of replicated particles in the first stage is $N_t = \sum_{m=1}^M N_t^{(m)}$, and in the second stage, $R_t = N - N_t$. The residual of the weight is obtained from

$$\hat{w}_t^{(m)} = w_t^{(m)} - \frac{N_t^{(m)}}{N}. \tag{12}$$

In the second stage, the particles are drawn according to the residual weights and by using multinomial resampling (or another random sampling method), where the probability for

selecting $x_t^{(m)}$ is proportional to the residual weight of that particle. Residual resampling has two loops taking on the order of $O(M) + O(R_t)$ time for computing.

The first stage represents a deterministic replication, and so the variation of the number of times a particle is resampled is only attributed to the second stage. Thus, the upper and lower limits of the number of times that the mth particle is resampled are $\lfloor N w_t^{(m)} \rfloor$ and $\lfloor N w_t^{(m)} \rfloor + R_t$, respectively, if the second stage is implemented using multinomial resampling. The code of the residual resampling method is given by code 4.

VARIATIONS OF TRADITIONAL RESAMPLING
The aforementioned four traditional methods are probably the best known and most used. They have been modified in various ways. For example, one of them removes the computationally expensive multinomial resampling part in residual resampling and implements resampling in a single loop (see code 4 in Table 2). The method is called *residual systematic resampling (RSR)* [12], [19]. RSR accumulates the fractional contributions of each particle in the searching sequence until it is large enough to generate a sample (which is equivalent to the accumulation idea used in systematic resampling). Then, no additional procedure is required for the residuals. Thus, one can have one iteration loop, and the complexity is of order $O(N)$.

If it is not mandatory to keep the particle size M constant at every time step and instead, the size is allowed to vary, we have simple ways of dealing with the particles in parallel and in one loop. We describe here two approaches. In the first approach, the number of replicated particles of $x_t^{(m)}$ is equal to $N_t^{(m)} = \lfloor N w_t^{(m)} \rfloor$ with probability $1 - p$ or equal to $N_t^{(m)} = \lfloor N w_t^{(m)} \rfloor + 1$ with probability p, where $p = N w_t^{(m)} - \lfloor N w_t^{(m)} \rfloor$. This method is called the *branch-kill* procedure [20] or *branching* [21]. In the second approach, $N_t^{(m)}$ is the nearest integer of $N w_t^{(m)}$, i.e., the rounding result of $N w_t^{(m)}$. We refer to this method as *rounding-copy* resampling [22]. Both methods (see code 5 in Table 2) require no additional operation and satisfy the unbiasedness condition but generate a varying sample size. The upper and lower limits of the number of replications of the mth particle in the branch-kill, rounding-copy, and RSR methods are all $\lfloor N w_t^{(m)} \rfloor$ and $\lfloor N w_t^{(m)} \rfloor + 1$, respectively.

[TABLE 3] A COMPARISON OF TRADITIONAL RESAMPLING METHODS (ALL THE RESAMPLED PARTICLES HAVE EQUAL WEIGHTS).

RESAMPLING METHODS	COMPUTATIONAL COMPLEXITY	ABLE TO KEEP CONSTANT SAMPLE SIZE	NUMBER OF RANDOM NUMBERS USED	NUMBER OF TIMES A PARTICLE IS REPLICATED	
				LOWER LIMIT	UPPER LIMIT
MULTINOMIAL	$O(NM)$ (OR $O(N \log M)$)	YES	N	0	N
RESIDUAL	$O(M) + O(R_t)$	YES	R_t	$\lfloor N w_t^{(m)} \rfloor$	$\lfloor N w_t^{(m)} \rfloor + R_t$
STRATIFIED	$O(N)$	YES	N	$\max(\lfloor N w_t^{(m)} \rfloor - 1, 0)$	$\lfloor N w_t^{(m)} \rfloor + 2$
SYSTEMATIC	$O(N)$	YES	1	$\lfloor N w_t^{(m)} \rfloor$	$\lfloor N w_t^{(m)} \rfloor + 1$
RESIDUAL SYSTEMATIC	$O(N)$	YES	1	$\lfloor N w_t^{(m)} \rfloor$	$\lfloor N w_t^{(m)} \rfloor + 1$
BRANCH-KILL	$O(N)$	NO	M	$\lfloor N w_t^{(m)} \rfloor$	$\lfloor N w_t^{(m)} \rfloor + 1$
ROUNDING-COPY	$O(N)$	NO	0	$\lfloor N w_t^{(m)} \rfloor$	$\lfloor N w_t^{(m)} \rfloor + 1$

A succinct comparison of the properties of the aforementioned methods is given in Table 3. The computational speeds of the methods are presented in [14] and [22].

COMPOUND SAMPLING

The resampling methods addressed so far are based on an approach where all the particles are sampled in the same way. This entails obtaining relatively similar resampling results. In all of the methods, the condition of unbiasedness is satisfied, and the resampled particles are equally weighted. In the following, we describe methods where resampling is realized without attempting to satisfy the conditions of unbiasedness and equal-weighting. This may entail risks of which the practitioner must be aware.

The compound sampling methods are based on grouping the particles by using predefined criteria prior to applying resampling. The groups are nonoverlapping, and they represent a partition of the whole particle population. The criterion for grouping is usually based on weight thresholds so that particles with similar weights are put in the same group, and then resampling is performed of each group in different ways. The reasons for applying compound resampling include decreasing the execution time and preserving particle diversity.

Compound resampling has its roots in stratified sampling. We will classify the methods based on whether the grouping is performed using a predefined threshold, which is either a constant or a function of the weights, or it is based on the particle values. We note that particles are often grouped when we implement parallel resampling. However, in parallel resampling, the groups are most commonly formed just by clustering index-neighboring particles.

THRESHOLD-/GROUPING-BASED RESAMPLING

In this category, particles are placed into different groups based on weight thresholds, and one uses different sampling strategies for each group to provide more flexibility for resampling. The threshold can be dynamic/adaptive or fixed, and there can be one or more thresholds.

Dynamic Threshold

The optimal resampling from [24] automatically sets a threshold value c_t, where c_t is a unique solution of

$$N = \sum_{m=1}^{M} \min\left(\frac{w_t^{(m)}}{c_t}, 1\right),\qquad(13)$$

where N is given, and $N < M$. All the particles whose weights are above this threshold are entirely preserved rather than replicated. Thus, there are no multiple copies of these particles in the final set of N particles. The other particles are resampled with a probability corresponding to their weights and assigned a weight c_t. We see that the resampled particles do not have equal weights (see code 6). The main advantage of the method is that, among the unbiased resampling methods, it is optimal in terms of minimizing the squared error-loss function

$$E\left(\sum_{m=1}^{M} (\tilde{w}_t^{(m)} - w_t^{(m)})^2\right),\qquad(14)$$

where $\tilde{w}_t^{(m)}$ is the new weight of $x_t^{(m)}$ when it is resampled; otherwise, $\tilde{w}_t^{(m)}$ is equal to zero. The method is appropriate for PF that uses increased number of propagated particles, and optimal resampling reduces the number to $N < M$. A disadvantage of the method is the need for calculating the value of c_t at each iteration. Also, the resampled particles may still suffer from degeneracy as the variance of the weights remains high.

Code 6: Optimal resampling.

$[\{\tilde{x}_t^{(n)}, \tilde{w}_t^{(n)}\}_{n=1}^{N}] = $ (Optimal)Resample $[\{x_t^{(m)}, w_t^{(m)}\}_{m=1}^{M}, N]$
Compute c_t satisfying (13)
$n = 0;\ h = 0$
FOR $m = 1: M$
   IF $w_t^{(m)} \geq c_t$
      $n = n + 1$
      $\tilde{x}_t^{(n)} = x_t^{(m)};\ \tilde{w}_t^{(n)} = w_t^{(m)}$
   ELSE
      $h = h + 1$
      $A^{(h)} = x_t^{(m)};\ B^{(h)} = w_t^{(m)}$
   END
END
$N_1 = n$
$[\{\tilde{x}_t^{(n)}\}_{n=N_1+1}^{N}] = $ (Stratified)Resample $[\{A^{(r)}, B^{(r)}\}_{r=1}^{h}, N - N_1]$
FOR $n = N_1 + 1: N$
   $\tilde{w}_t^{(n)} = c_t$
END

There are similarities between optimal resampling, rejection control (RC) [17] and partial RC (PRC) [11]. In RC, a control threshold c_t is computed, which may be a quantity given in advance, e.g., the median or a quantile of the weights, and the mth particle is accepted with a probability

$$p = \min\left(1, \frac{w_t^{(m)}}{c_t}\right).\qquad(15)$$

In PRC, the particles with weights that are greater than or equal to c_t are automatically accepted, whereas other particles are accepted with probability p. This can be viewed as a combination of the rejection method and importance sampling. An accepted particle $x_t^{(m)}$ is reweighted with a new weight $\max(c_t, w_t^{(m)})$, and the rejected particles are replaced by particles regenerated from particles from previous time instances. These two forms of RC differ primarily in how far one goes back to regenerate particles. The RC does it to the earliest time, $t = 0$, while the PRC only regenerates particles from time instant $t - 1$ so that one saves on computations. These methods cannot be considered as candidates for real-time implementation because they have nondeterministic execution time and large memory requirements.

Fixed Threshold

A partial deterministic reallocation approach is proposed in [11] based on a fixed threshold, say $1/N$, where N is the desired sample size. The mth particle with a weight larger than $1/N$, is replicated $\lfloor Nw_t^{(m)} \rfloor$ (or $\lfloor Nw_t^{(m)} \rfloor + 1$) times, and the weights

after resampling are $w_t^{(m)}/\lfloor Nw_t^{(m)}\rfloor$ (or $w_t^{(m)}/(\lfloor Nw_t^{(m)}\rfloor+1)$) (this is referred to as *particle splitting* in the article). The mth particle with weight smaller than $1/N$ is sampled with probability $Nw_t^{(m)}$ and is weighted as $1/N$ (see code 7). The resampled particles are not equally weighted, and, notably, their weight sum is not one (a normalization step is additionally required). The sampling method is biased.

Code 7: Reallocation resampling.

$$\left[\{\tilde{x}_t^{(n)}, \tilde{w}_t^{(n)}\}_{n=1}^{N^*}\right] = (\text{Reallocation})\text{Resample}\left[\{x_t^{(m)}, w_t^{(m)}\}_{m=1}^{M}, N\right]$$

$n = 0$
FOR $m = 1: M$
  IF $w_t^{(m)} \geq 1/N$
    $N_t^{(m)} = \text{Floor}\,(N \times w_t^{(m)})$ (or $N_t^{(m)} = \text{Floor}\,(N \times w_t^{(m)})+1$)
    FOR $h = 1: N_t^{(m)}$
      $n = n + 1$
      $\tilde{x}_t^{(n)} = x_t^{(m)};\ \tilde{w}_t^{(n)} = w_t^{(m)}/N_t^{(m)}$
    END
  ELSE
    $u \sim U(0,\ 1/N]$
    IF $w_t^{(m)} \geq u$
      $n = n + 1$
      $\tilde{x}_t^{(n)} = x_t^{(m)};\ \tilde{w}_t^{(n)} = 1/N$
    END
  END
END
$N^* = n$

To reduce the complexity of hardware realization and the required power consumption, resampling is performed on only some of the particles. This is the idea behind partial resampling (PR) [12]. PR consists of two steps: in the first, the particles are classified as moderate, negligible, or dominating; and in the second, different resampling strategies are applied to each group of particles. In [17], three different resampling functions are proposed for determining which particles are resampled/discarded and how to allocate the weights, and, to that end, two thresholds are used. To further increase the processing speed, the classification of the particles can be overlapped with the weight computation (overlapped PR).

Grouping strategies may also be applied to alleviate the impoverishment that is practically inevitable in all the resampling methods presented so far. For example, the double systematic resampling approach from [25] partitions the particles into two groups, in low- and high-weighted particles and where the number of particles to be resampled from each group is specified. In this way, the low-weighted particles are resampled independently from the group of high-weighted particles. Obviously, this may lead to biased sampling.

There are several other methods from the literature that are built on similar ideas. Before their adoption, however, their features need to be better understood and firm guidelines for use provided.

RESAMPLING THAT TAKES INTO ACCOUNT PARTICLE VALUES

All of the above resampling methods are based on the particle weights. A possible way to improve resampling is to exploit the particle values (the state information they contain) during resampling. The particle distribution in the state space exhibits their diversity and, therefore, is the key for monitoring impoverishment. In this category of resampling methods, the particles are grouped with respect to not only their weights but also their values.

Particles in close proximity in values may represent a similar state, and, thus, they can be merged to reduce the number of particles with different values. This is the basis of deterministic resampling [16], which replaces the second stage of residual resampling by merging particles using their residuals. This aims at preserving the diversity of the particles so that no particles are discarded at all, which is helpful when the number of particles is small. Contrary to particle merging, particle splitting replaces a particle with a large weight (larger than a threshold) with a set of particles with the same values and whose sum of weights is equal to the original weight (see Figure 2). Particle merging is implemented before the weight-updating step to reduce the sample size to save computation, while particle splitting is applied after weight updating as an alternative to resampling to reduce the weight variance and to recover the sample size [23]. A main disadvantage of these methods is that the dimensionality-free property of resampling is destroyed.
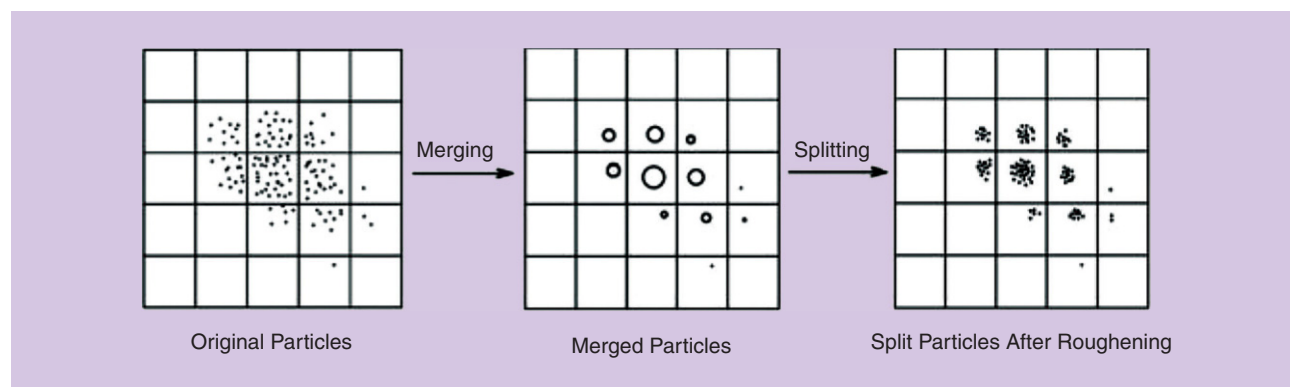
[FIG2] Particle merging and splitting in a two-dimensional state space. The size of the circles represents the weight of particles.

Resampling Methods	Grouping and Sampling Methods Adopted				Benefits Reported	Disadvantages
	Sampling Methods (Particles with Low Weights)	Thresholds (Defining Strata)		Sampling Methods (Particles with High Weights)		
Optimal Resampling [24]	Stratified Sampling	Dynamic, c_t Where c_t Is a Unique Solution of (13)		Preserved	Minimized Squared Error-Loss Function	Additional Computation of (13); The Sample Size Decreases
(Partial) RC [17]	Replaces Particles by Regenerating Them from Previous Particles	Dynamic, c_t Where c_t May Be a Quantity Given in Advance or the Median or a Quantile of the Weights		Preserved and Reweighted with a New Weight (Reject Sampling)	Advantageous in Dealing with Sudden Changes in the Dynamic System	Increased Computational Intensity
Reallocation [11]	Rejection Sampling	$1/N$		Split	Reduced Resampling Variance (As Compared with Traditional Resampling Methods)	Varying Sample Size
PR [12]	(Small Weight) Preserved or Discarded	Fixed Threshold 1	(Moderate Weight) Preserved and Reweighted with Equal Weights	Fixed Threshold 2 / (Significant Weight) Preserved and Reweighted with Equal or Different Weights	Reduced Computational Complexity	Thresholds Need to Be Carefully Specified
Deterministic Resampling [16]	Merged	$1/N$		Integer Parts ($1/N$) Are Replicated and Residuals Are Merged	Preserved Diversity of Particles (Without Discarding Any Particle)	Sensitive to the Dimensionality of the State
Double Systematic Resampling [25]	Systematic Resampling	$1/N$		Systematic Resampling	Increased Resampling Flexibility (Higher Probability for Drawing Small-Weighted Particles)	Two Independent Weight Normalization and Resampling

[FIG3] A comparison of several compound sampling methods.

COMPARISON OF SEVERAL COMPOUND SAMPLING METHODS

We emphasize that the core idea of compound sampling is to deal with particles from different groups differently. In Figure 3, we provide a succinct overview of different compound sampling methods, where the following terms for particles are used:

- *Discarded*: Particles that are discarded and not resampled.
- *Preserved*: Particles that are preserved with their weights being kept unless otherwise stated.
- *Merged*: Particles (in a specified space) that are merged to one particle with a state value equal to the weighted mean of the states of merged particles and with a weight equal to the sum of the original weights of the merged particles.
- *Split*: A particle that is divided into several copies. The split copies have the same state and weight, and their sum of weights is equal to the weight of the original particle.

- *Replicated*: A particle that is replicated, and each copy has the same value as the original particle as well as the same weight, unless otherwise stated.

Furthermore, in Table 4, we provide features of some of the compound sampling methods. They include optimal resampling [24], (partial) RC [17], reallocation [11], partial resampling [12], deterministic resampling [16], and double systematic resampling [25].

SPECIAL STRATEGIES

As previously mentioned, the key to combating degeneracy while avoiding impoverishment by resampling is the introduction of a compromise between concentration (the replication of large-weighted particles) and diversification (the discarding of negligible particles). To that end, several strategies have been proposed, including modified resampling, variable-size resampling, and roughening.

[TABLE 4] THE SPECIAL PROPERTIES OF SEVERAL COMPOUND RESAMPLING METHODS.

BIASED SAMPLING	UNEQUAL WEIGHTED	VARYING NUMBER OF PARTICLES	STATE CONSIDERED	UNIQUE/OUTSTANDING HIGHLIGHTS
[11], [12], [17], [25]	[11], [12], [16], [17], [24],	[11], [16], [24]	[16]	[11]: PARTICLE SPLITTING IS APPLIED; THE WEIGHT SUM OF THE RESAMPLED PARTICLES IS NOT ONE [12]: TWO (OR MORE) THRESHOLDS ARE USED [16]: AVOIDS DISCARDING ANY PARTICLE [17]: PREVIOUS PARTICLES ARE CONSIDERED; REJECT METHOD IS USED [24]: MINIMIZE THE SQUARED ERROR LOSS FUNCTION [25]: LOW-WEIGHTED PARTICLES ARE PROTECTED

In brief, the idea behind modified resampling is to draw particles from a distribution derived from the weights of the particles. The variable-size resampling, as the name suggests, provides a different number of samples with time so that a predefined criterion is satisfied (e.g., to reduce computational requirements, one uses a smaller number rather than a large number of particles, or to improve accuracy of tracking, one draws a larger number of particles). Roughening entails perturbing the locations of the resampled particles so that we reduce impoverishment, and, thus, it is performed once the resampling is completed. Next, we describe each of the strategies in more detail.

MODIFIED RESAMPLING

In generalized resampling [11], particles are resampled according to the probabilities $p_t^{(m)}$. The latter are usually equal to the particle weights $w_t^{(m)}$, but more generally, one can draw particles with probabilities that are functions of the weights, i.e.,

$$p_t^{(m)} \propto (w_t^{(m)})^{\alpha}, \tag{16}$$

where $\alpha > 0$. When $0 < \alpha < 1$, the low-weighted particles get boosted, and the large-weighted particles have suppressed probabilities, and, thereby, the particle diversity improves. By contrast, $\alpha > 1$ entails increased preference for higher-weighted particles.

Knowledge about the next observation before resampling can be implemented via auxiliary weights [18]. In that way, the particles that are likely to have higher likelihoods have a better chance of surviving. There, the step of generating the auxiliary variable, which represents the fitness of the particle, can be viewed as a resampling step that takes into account both the immediate future and present states when carrying out selection. It is an appealing idea to fuse the information from the newest observations with the current weights while making a decision on the selection of particles.

VARIABLE-SIZE RESAMPLING

The use of a constant number of particles is not always the best choice because the complexity of the distributions of interest can vary drastically over time. In obtaining the number of particles that is both efficient and sufficient for approximating the distribution of interest, the underlying idea is to choose a small number of particles if the distribution is focused on a small part of the state space and to adopt a large number of particles if the distribution is much more spread out. This is the core rationale of the Kullback–Leibler divergence (KLD)-sampling approach [26], which determines the needed number of particles based on the KLD between the sample-based maximum likelihood estimate and a distribution of interest.

The required number N of particles can be determined so that, with the probability $1 - \rho$, the KLD between the sample-based maximum likelihood estimate of a desired distribution and that distribution is less than a prespecified error bound threshold ε. In [26], it is found that

$$N = \frac{1}{2\varepsilon} q, \tag{17}$$

where

$$q = F^{-1}(1 - \rho), \tag{18}$$

with $F^{-1}(\cdot)$ being the inverse of the cumulative chi-squared distribution with $k - 1$ degrees of freedom, and k the number of bins [nonoverlapping (multi)dimensional intervals] used for sorting the particles. The value of N in (17) could be approximately computed [26]. In practice, the number of particles for resampling may be hard-limited to be not less than a minimum threshold.

Ideally, one would want to use as a desired distribution the posterior of the state. In [26], the posterior is approximated by the predictive distribution. Theoretically, it is more rigorous and flexible to apply (17) during resampling than in sampling, which leads to KLD-resampling [27]. There, the particles are resampled one by one until the required amount given by (17) is reached. Obviously, the disadvantage of the KLD-based method is that the particles need to be sorted out in bins defined on the state space, which can be very computationally costly.

ROUGHENING STRATEGIES

In obtaining an optimal set of particles, instead of designing the optimal proposal distribution, one can employ compensation. If impoverishment has already occurred after resampling, one approach to reducing it is to spread the overcentralized particles by roughening or jittering their values. This simply means that we add random noise (roughening noise) to the resampled particles. The roughening noise is normally Gaussian with zero mean and constant covariance. In [1], it is suggested that the noise covariance is diagonal with a standard deviation for a particular state component given by $\sigma = K D N^{-1/d_x}$, where D is the difference between the maximum and minimum values of the state component, K is a positive tuning constant chosen by the user, N is the number of particles, and d_x is the dimension of the state. The roughening may be applied 1) only at selected steps, 2) only to selected particles that are resampled from the same particle, and

[TABLE 5] PARALLEL PLATFORMS FOR PF IMPLEMENTATION.

PLATFORM	ADVANTAGES	DISADVANTAGES	APPLICATION
VLSI [31]	■ CUSTOMIZED ARCHITECTURE FOR SPECIFIC APPLICATION ■ VERY HIGH PROCESSING SPEED CAN BE ACHIEVED	■ VERY COSTLY ■ VERY LONG DESIGN CYCLES	■ NOT USED YET—SUITABLE FOR EXTREMELY LARGE QUANTITIES
FPGA [32]	■ CUSTOMIZED ARCHITECTURE FOR SPECIFIC APPLICATION	■ LIMITED FLOATING-POINT CAPABILITIES ■ LONG DESIGN CYCLE AND NEED TO KNOW HARDWARE AND HARDWARE DESIGN LANGUAGES	■ PROTOTYPING PLATFORM FOR REAL-TIME SYSTEMS
MULTICORE CENTRAL PROCESSING UNIT [33], [34]	■ COURSE-GRAINED PARALLELISM ON TWO TO EIGHT CORES. RELATIVELY EASY TO ACCESS AND TO PROGRAM	■ LARGE AMOUNT OF SEQUENTIAL PROCESSING	■ ACCELERATING SIMULATIONS
GP-GPU [35]–[37]	■ FINE-GRAINED PARALLELISM CAN BE ACHIEVED ■ AVAILABLE ON ALMOST EVERY COMPUTER ■ EASY TO PROGRAM	■ KNOWLEDGE OF GPU PROGRAMMING LANGUAGES SUCH AS CUDA OR OPENCL	■ ACCELERATING SIMULATIONS AND SOME HIGH-PERFORMANCE REAL-TIME APPLICATIONS
EMBEDDED GPU	■ PROGRAM IN THE SAME WAYS AS GP-GPU ■ CANDIDATE FOR PF FOR EMBEDDED SYSTEMS	■ LESS PARALLELISM AVAILABLE THAN IN GP-GPU ■ KNOWLEDGE OF GPU PROGRAMMING LANGUAGES	■ NEW PLATFORMS FOR EMBEDDED SYSTEMS SUITABLE TO REAL-TIME PERFORMANCE
CLUSTER AND MULTICOMPUTERS [33], [39]	■ COMPLEX PARTICLE FILTERS CAN BE SIGNIFICANTLY ACCELERATED	■ KNOWLEDGE OF PARALLEL PROGRAMMING LANGUAGES SUCH AS OPENMP OR MESSAGE PASSING INFERENCE	■ ACCELERATING SIMULATIONS

3) only in a few dimensions of the state space. Similar ideas for diversification can be implemented by using Markov chain Monte Carlo (MCMC) sampling. The so-called resample-move algorithm from [28] has a move step after the resampling step based on MCMC sampling. The move step performs one or more iterations on each of the resampled particles, thereby rejuvenating the diversity of particles.

An important alternative to the roughening strategy is the use of Gaussian kernels. With the kernels, we smooth the posterior density by convolving each particle with a diffusion kernel. This is also known as *regularization*, and it amounts to replacing the discrete distribution defined by the particles and their weights with a continuous approximation [29]. The resampling step is then changed to simulations from the continuous distribution, which generates a new particle set. All of these approaches produce diversified copies of the same particles.

PARALLEL PROCESSING
Despite its successful applications, PF often suffers from a heavy computational cost, especially when the dimension of the state-space model is high and the number of used particles has to be large. However, the execution of PF can be accelerated through parallelization for both offline and real-time processing. An important area where parallelized PF also takes place is in signal processing over networks, where PF operates in parallel at the nodes of the network [30]. There, the problems of PF are of a different nature than the ones we address here. In this article, we focus on describing the parallelization of sequential PF and the resampling algorithm.

> **PF OFTEN SUFFERS FROM A HEAVY COMPUTATIONAL COST, ESPECIALLY WHEN THE DIMENSION OF THE STATE-SPACE MODEL IS HIGH AND THE NUMBER OF USED PARTICLES HAS TO BE LARGE.**

PARALLEL PROCESSING ARCHITECTURES FOR PF
Parallel hardware devices where PF has been implemented and reported in the literature include custom very large-scale integration (VLSI) chips [31], field-programmable gate arrays (FPGAs) [32], multicore processors [33], [34], GP-GPUs [35]–[38], and computer clusters/multicomputers [33], [39]. The advantages, disadvantages, and potential applications of these platforms for PF implementation are shown in Table 5. Hardware implementations on VLSI and FPGA result in high-speed implementation since every mathematical function is customized and implemented in hardware. However, VLSI design is very expensive, and, to the best of our knowledge, there are no commercial chips with implemented PF yet. The number of floating point units that can be executed in parallel on an FPGA is smaller than on state-of-the-art GP-GPUs. Also, floating-to-fixed-point conversion is very complicated, and particle filters require a large number of bits in fixed-point representation. In addition, programming of GPUs is simpler than designing hardware for FPGAs. Multicore platforms and GP-GPUs are now available on almost every computer so that lately we have been witnessing more parallel PF implementations on GP-GPUs.

Recently, new chips with embedded multiple processors and GP-GPUs have appeared from various companies. They support the same programming models as GP-GPUs, including CUDA and Open Computing Language (OpenCL). They are intended for high-performance embedded applications and have much lower power consumption than GP systems. Furthermore, many-core chips that have more than 16 processors on a single chip have appeared and have started to be used for high-performance

embedded applications. As far as we know, PF has not yet been implemented on these embedded platforms.

CHALLENGES

The particle generation and weight updating are computationally the most intensive steps and can be implemented in parallel if parallel processing hardware is available. Resampling, normalization, and computing estimates are inherently sequential, and they limit the speedup that can be achieved with parallel processing. An additional step, called *particle allocation*, is usually needed for efficient parallel implementation. We describe this step with an example. Assume that parallel processing is performed using a generic architecture with three PEs. Let each PE perform sampling and weight computation of an equal number of particles equal to $M/3$, where $M/3$ is an integer. After resampling, some particles are replicated, and some are discarded. Now consider a situation where all the particles in PE1 and PE3 are discarded, and the majority of the particles in PE2 are replicated. When the next observation comes, there will be no particles in PE1 and PE3 for processing, and almost all the processing will occur sequentially in PE2, resulting in low PE utilization and load imbalance. To avoid this situation, particles from PE2 need to be redistributed (allocated) to the other PEs so that all the PEs have an equal amount of particles before the next step. The particle allocation can be done by sending particles and their weights in message-passing architectures or by reallocating particles (or their indexes) in the global memory in shared-memory architectures. Particle allocation requires additional time, and, consequently, it affects the speedup.

Major challenges for the parallel implementation of PF include:

■ There might be computationally complex operations in the sampling and weight-computation steps, including random number generation. Some of these operations might not be supported by all processing architectures.

■ Operations such as normalization, computation of output estimates, and computation of cumulative sums in some resampling algorithms are sequential and require particles from all PEs.

■ Particle allocation requires frequent communication with the global shared memory in shared-memory systems, which is a bottleneck. Similarly, the communication between the PEs in message-passing systems consumes significant time.

■ Platform-specific challenges are related to the fact that every platform requires modifications or implementations that are platform specific. For example, design of a hardware random number generator is required in FPGA platforms, while an efficient parallel software algorithm for random number generation is needed for GPUs.

PERFORMANCE METRICS AND TERMINOLOGY

The major objective of parallel implementation is to reduce the execution time, and the most commonly used metric is speedup [34],

[35]. We define the execution time of particle filters as the time they need to process one observation [31], [32], [36]. The speedup is defined as the ratio of the execution time of the best possible serial algorithm (on a single processor) to the execution time of the chosen algorithm on a parallel system based on multiple processors. The efficiency is defined as the speedup divided by the number of processors [33]. System utilization indicates the percentage of resources that was kept busy during the execution of a parallel program. Communication overhead is defined as the percentage of time spent on interprocess communication and all noncomputing operations [33]. The degree of parallelism is a measure of the numbers of threads of computation that can be carried out simultaneously. Data dependencies are the result of precedence constraints between operations, and they prevent concurrent execution. For example, the degree of parallelism for the weight-computation step is equal to the number of particles. A number of papers have been devoted to increasing the degree of parallelism and system utilization during resampling to decrease execution time.

> **IF IMPOVERISHMENT HAS ALREADY OCCURRED AFTER RESAMPLING, ONE APPROACH TO REDUCING IT IS TO SPREAD THE OVERCENTRALIZED PARTICLES BY ROUGHENING OR JITTERING THEIR VALUES.**

Other types of performance metrics are related to tracking the performance of the parallel particle filter. Researchers mainly rely on examples to monitor the change of the mean square error [32], [34] or the effective sample size [33], [40] of the parallel versus sequential implementation of PF. Theoretical work on the performance or loss of performance of PF due to parallel implementation is still missing. One attempt is made in [45], where the analysis of the sample variance of the weights, the distortion of the probability measure, and the variance of estimators of a distributed PF are given. In [33], the mean and standard deviation of the log likelihood are calculated to validate that the resampling procedure does not affect the performance of the particle filter.

DISTRIBUTED RESAMPLING

Much work has been focused on parallelization of traditional resampling, referred to as *distributed resampling*. Two main classes of approaches have been applied: 1) the algorithm is modified in a way that the result of resampling is not changed in comparison with the sequential algorithm, and 2) the algorithm is modified, and it has been shown that, mainly through simulations, its parallel implementation provides similar results to the original sequential resampling. The majority of the algorithmic modifications is performed to reduce the communication overhead mainly during the particle allocation step.

An example of an algorithm that does not change the result of standard resampling is distributed resampling with proportional allocation [32]. Here, the resampling step is divided into two steps. In the first step, each PE is considered as a particle whose weight is the cumulative weight of all the particles of this PE. The resampling is performed only on the cumulative weights to determine how many particles each PE needs to generate. Next, local resampling is performed in parallel in each PE. In this way, the degree of

parallelism for the resampling step is equal to the number of PEs. Subsequently, the particle allocation step is performed, in which the excess of particles after resampling from some PEs is communicated or made available to the PEs with the deficiency of particles. Central (sequential) processing is still needed for the first step of resampling, normalization, and computing the output estimate. A number of variations of this method have been devised where load unbalance and the amount of central processing are reduced.

The load balancing solution for the GP-GPU platform has been proposed in [36] using CUDA. All the particles are stored in a global GP-GPU memory, which is a shared memory, and writing to it is slow. During the particle-allocation step, the same replicated particles are written many times to the memory. After the particle allocation, each set of particles from the global memory is apportioned to different parallel threads. The optimization performed here is architecture (GP-GPU) specific.

> **SERIAL COMPUTATION AND GLOBAL COMMUNICATION ARE INEVITABLE IN WEIGHT NORMALIZATION, AND, THEREFORE, ALTERNATIVES THAT ARE FREE OF NORMALIZATION BECOME ATTRACTIVE.**

The aforementioned methods require some central (sequential) processing to perform resampling. To further improve the speed of resampling, it is desirable to perform local particle allocation without having a central unit. Localized particle allocation produces different results than traditional resampling algorithms, but it reduces communication overhead, making it concurrent. In localized resampling, only a subset of the particles is exchanged between neighboring PEs/threads, and the exchange is preferably performed in parallel. The resampling and all the other operations are also performed in parallel. Various solutions have been proposed that depend on different parameters, including the number of PEs, the number of particles to be exchanged, the types of particles to be exchanged, and the selection of PEs where the particles are sent. Resampling with nonproportional allocation (RNA) [32] allows for each PE to perform local resampling and to exchange some fixed number of particles with one neighboring PE. The platform where PF was implemented in [32] was an FPGA. RNA-based implementation of parallel resampling is carried out in one of the first implementations of PF on GP-GPU [37]. In other implementations, one or several of the largest particles are exchanged with neighboring PEs. As an example, [38] presents a GP-GPU implementation where local sorting is performed to order the particles that are being exchanged. In [40], a method is proposed for finding the optimal portion of particles for exchange between adjacent PEs and for achieving the mixing of the posterior distributions among the adjacent PEs, so as to preserve particle diversity.

NORMALIZATION-FREE RESAMPLING

Serial computation and global communication are inevitable in weight normalization, and, therefore, alternatives that are free of normalization become attractive. In this category, resampling is performed free of weight normalization, differing from all the resampling methods presented so far. In brief, the particles are (re-)sampled based on their relative magnitude of the (nonnormalized)

weights, e.g., the ratios between the weights globally, or the absolute weight comparison between local neighboring (two or three) particles. The former is globally unbiased sampling, while the latter is local sampling that is almost surely biased.

Ratios Between Weights

Metropolis [38] and independent Metropolis–Hastings sampling [39] require only ratios between weights that do not need to be normalized and therefore threads can process in parallel, see code 8, for example. The solution presented in [38] addresses GP-GPU implementation where the amount of parallel computational resources is abundant, and it stresses that even though Metropolis sampling is more computationally intensive than traditional resampling algorithms, it is as fast on a GP-GPU because there are no dependent operations on the weights. There, sampling from $U(1, \ldots M)$ returns a value randomly selected from the set $\{1, \ldots, M\}$.

Code 8 is an iterative process of sampling based on constructing a Markov chain. More specifically, it uses a Metropolis–Hastings move step for searching in a particle set for a particle with a large weight to replace the current particle. The depth of the search (burn-in) is denoted by B. It is desirable that the number of times a particle is sampled is proportional to its weight. As in most MCMC algorithms, deeper searching will provide better results (closer to the desired distribution).

The selection of B is a tradeoff between speed and reliability. While runtime is reduced with fewer steps, the sample will be biased if B is too small to ensure convergence. Similarly, a number of additional particles for burn-in (the time during which the Markov chain is in a transient state) is needed. The particles that are generated in the burn-in period are discarded. A disadvantage of this method is that, in most cases, it is difficult to estimate the required burn-in period.

Code 8: Metropolis resampling.

$[\{\tilde{x}_t^{(m)}, \tilde{w}_t^{(m)}\}_{m=1}^M] = $ (Metropolis)Resample $[\{x_t^{(m)}, w_t^{(m)}\}_{m=1}^M, B]$
FOR $m = 1:M$ (can be in parallel)
  $k = m$
  FOR $n = 1:B$
    $u \sim U(0,1);\; l \sim U\{1, \ldots, M\}$
    IF $u \leq w_t^{(l)}/w_t^{(k)}$
      $k = l$
    END
  END
  $\tilde{x}_t^{(m)} = x_t^{(k)};\; \tilde{w}_t^{(m)} = 1/M$
END

Local Neighbor-Comparison Methods

Now, we consider more straightforward local methods that include local Monte Carlo [10] and Local Selection (LS) methods [41]. The

local methods consist of randomly sampling a particle from a small set of successive particles, where the sampling is based only on the weights of the particles in the set. The resampled particles preserve their weights. Assume, for simplicity, that the number of PEs K is equal to the number of particles M (if $K < M$, particles allocated to the same PE are processed serially), and the mth PE contains a particle $\{x_t^{(m)}, w_t^{(m)}\}$ in its memory, where $w_t^{(m)}$ is not normalized. Before the parallel resampling, a communication step is required, where the mth PE transmits its particle to its two neighboring PEs. Then, each PE contains three particles, and the LS resampling can be realized as by code 9.

For consistency of description, note that $w_t^{(0)}$ refers to $w_t^{(M)}$, while $w_t^{(M+1)}$ refers to $w_t^{(1)}$.

Obviously, the number of replications of each particle in LS is very limited in local groups and is maximally three. This can cause a biased result, especially if high-weighted particles are next to each other. Then, some of them will be discarded and the performance will be degraded.

> [**THE RESAMPLING PROCESS REPRESENTS A BOTTLENECK IN PARALLEL IMPLEMENTATIONS.**]

Code 9: Local selection resampling.

$[\{\tilde{x}_t^{(m)}, \tilde{w}_t^{(m)}\}_{m=1}^M] = (\text{LS})\text{Resample }[\{x_t^{(m)}, w_t^{(m)}\}_{m=1}^M]$
FOR $m = 1: M$ (in parallel)
$\quad W = w_t^{(m-1)} + w_t^{(m)} + w_t^{(m+1)}$
$\quad T_1 = w_t^{(m-1)}/W$
$\quad T_2 = (w_t^{(m-1)} + w_t^{(m)})/W$
$\quad u \sim U(0,1]$
$\quad$ IF $u \leq T_1$
$\quad\quad \tilde{x}_t^{(m)} = x_t^{(m-1)}; \ \tilde{w}_t^{(m)} = w_t^{(m-1)}$
$\quad$ ELSE IF $T_1 < u \leq T_2$
$\quad\quad \tilde{x}_t^{(m)} = x_t^{(m)}; \ \tilde{w}_t^{(m)} = w_t^{(m)}$
$\quad$ ELSE
$\quad\quad \tilde{x}_t^{(m)} = x_t^{(m+1)}; \ \tilde{w}_t^{(m)} = w_t^{(m+1)}$
$\quad$ END
END

We note that although weight normalization is avoided in these normalization-free resampling approaches, it is still necessary for computing the filtering estimates.

OTHER TYPES OF PARALLELIZATION OF PF

We reiterate that the resampling process represents a bottleneck in parallel implementations. For this reason, there have been efforts to develop PF methods that do not require resampling [see first item below] and overlap the operations of multiple PF or the steps of the same PF so that all the operations are sequential but are executed concurrently (second and fourth items below). More specifically, these efforts include the following:

1) Removing the resampling step from PF, such as in nonresampling PF detector [42] and Gaussian PF (GPF) [43], which are resampling-free.

2) Running several PFs independently on separate processors or a set of agents in a distributed network [30]. Some or all of the agents perform local PF and interact with other agents to calculate a global state estimate. These decentralized agent networks do not include a central unit.

3) Decomposing the state into two parts and considering the filtering problem as two nested subproblems. These two problems are then handled by separate PFs. This is also referred to as *decentralized PF* [44].

4) Pipelining the sampling and resampling: when a particle is resampled, it processes the sampling (particle propagation and weight updating) ahead that does not need to wait for other particles, i.e., the sampling in the next iteration will be produced before the resampling is finished.

Before concluding the subsection on parallel processing, we list some topics for further research on it. They include:

■ Applying PF to real-world applications by implementing them on embedded multicore, embedded GPUs, and many-core chips and guaranteeing real-time performance (for example, current GPUs do not incorporate hard real-time features).

■ Deriving optimization criteria that allow for evaluation of the quality of the localized distributed resampling algorithms versus sequential traditional resampling algorithms. Evaluating the practical benefit of parallel processing in terms of not only computing speed but also filtering accuracy. Some initial comparisons of parallel implementations of resampling that offer theoretical analysis and simulations are available. For example, analysis of resampling via RNA and LS in terms of reduction of the sample variance of the weights, the distortion of the probability measure, and the variance of estimators is given in [45].

■ Devising new parallel algorithms for specific architectures—for example, communication is very expensive in GP-GPUs, and therefore, one research direction is deriving algorithms with reduced communication between the PEs that can even be more computationally intensive.

FREQUENCY OF RESAMPLING

The benefits of resampling are accompanied with potential side effects such as sample impoverishment and prevention of parallel processing as explained above. Thus, resampling should only be applied when necessary and therefore it is important to have a method for determining how frequently or when to implement resampling. Two schedules for resampling have been proposed: deterministic and adaptive. In a deterministic schedule, one does resampling at fixed times $t_1, t_2, \ldots$, where t_i is often chosen to be $i \times t_1$. In an adaptive schedule, the times at which resampling occurs are selected by checking a criterion that assesses the quality of the current weights. In that case, whenever the criterion for quality is below a given threshold, the resampling step is triggered. This tends to produce better performance of PF than deterministic scheduling due to its flexibility.

As a key to adaptive schedules, the criterion for implementing resampling is usually based on the variation of the weights,

which reflects the degree of weight degeneracy. One such criterion is the effective sample size (ESS) N_{eff} defined by [46]

$$N_{\text{eff},t} = \frac{M}{1 + \text{Var}(w_t^*)}, \qquad (19)$$

where w_t^* is a nonnormalized weight, and the variance is computed with respect to the sampling distribution. Typically, obtaining N_{eff} from (19) is impossible [46]. Instead, one may employ the rule of thumb estimate given by

$$N_{\text{eff},t} = \left(\sum_{m=1}^{M} (w_t^{(m)})^2 \right)^{-1}. \qquad (20)$$

An application of the Cauchy–Schwarz inequality leads to the (intuitive) conclusion that $N_{\text{eff},t} \le M$. It is also clear that $1 \le N_{\text{eff},t} \le M$. Resampling occurs when the ESS falls below a selected threshold, γ_r. If γ_r is set to $\gamma_r = 0$, resampling never takes place, and $\gamma_r = M$, meaning that resampling occurs at every time step. Several different criteria to calculate the ESS have also been proposed; see, e.g., [8], [47], and [48]. It is necessary to note that criteria based on (20) are primarily used in tracking low-dimensional states. In high-dimensional problems, other metrics may be more appropriate.

CONCLUSIONS

In this article, the state of the art of resampling methods was reviewed. The methods were classified and their properties were compared in the framework of the proposed classifications. The emphasis in the article was on the classification and qualitative descriptions of the algorithms. The intention was to provide guidelines to practitioners and researchers.

Some final comments:

■ The resampling methods can hardly output much different results if they satisfy the unbiasedness condition, preserve a constant number of particles, and equally weight the resampled particles.

■ If these restrictions are removed, some benefits may be obtained, e.g., adaptive adjustment of the number of particles, preservation of particle diversity, and alleviation of impoverishment.

■ Compound sampling and special strategies, such as modified resampling and variable-size resampling, and compensations after resampling provide more flexibility. They balance the necessity for diversity and the need for concentration that lies in the center of sample degeneracy and impoverishment. They may offer better approximation and benefits in practice, but often at the price of higher computational costs.

■ The normalization of weights and load imbalance of particles after resampling are main barriers for parallelization of resampling. Ways to carry out normalization, to output filter estimate, and to manage communication between PEs distinguish existing parallel PF algorithms.

■ An issue that we have not discussed is the theoretical effects of resampling on the convergence of the PF estimates. The resampling step is crucial for uniform convergence results, and some recent theoretical results on this and related to deterministic and random resampling can be found in [49].

■ The research on the resampling of particle filters is going in multiple directions:

• implementation specific
 – simplifying resampling algorithms for real-time implementation
 – adjustment of the algorithms to specific hardware/processing architectures
 – parallelization of the resampling
 – application and acceleration of resampling to non–PF-based problems such as importance sampling and the forward-backward algorithm [50]

• theoretical analysis
 – analysis of the features of resampling algorithms without considering them as a part of PF
 – analysis of the effects of different resampling algorithms on the PF convergence and accuracy of tracking.

AUTHORS

Tiancheng Li (t.c.li@mail.nwpu.edu.cn) received his bachelor's degree in mechanical and electrical engineering, with a minor in business administration, from Harbin Engineering University, China, in 2008, and his Ph.D. degree in electrical and electronic engineering from London South Bank University, United Kingdom, in 2013. He is a currently a Ph.D. candidate in mechatronics engineering at Northwestern Polytechnical University, China, and a postdoctoral associate with the Bioinformatic, Intelligent Systems, and Educational Technology Group, University of Salamanca, Spain. His research interests are in the general area of statistical signal processing, information fusion, system modeling with a particular focus on particle filtering, finite set statistics, and multiple object tracking.

Miodrag Bolić (mbolic@eecs.uottawa.ca) received his B.Sc. and M.Sc. degrees in electrical engineering from the University of Belgrade, Serbia, in 1996 and 2001, respectively, and his Ph.D. degree in electrical engineering from Stony Brook University, New York, in 2004. Since 2004, he has been with the University of Ottawa, Canada, where he is an associate professor with the School of Electrical Engineering and Computer Science. His research interests include signal processing architectures, hardware/software accelerators, biomedical signal processing and instrumentation, and the Internet of Things. He is director of the Computer Architecture, Radio-Frequency Identification, and Medical Devices Research Groups at the University of Ottawa. He is a Senior Member of the IEEE.

Petar M. Djurić (petar.djuric@stonybrook.edu) received his B.S. and M.S. degrees in electrical engineering from the

University of Belgrade, Serbia, and his Ph.D. degree in electrical engineering from the University of Rhode Island. He is currently a professor in the Department of Electrical and Computer Engineering at Stony Brook University, New York. His research has been in signal and information processing, with an emphasis on Monte Carlo-based methods, signal processing over networks and applications in wireless sensor networks, and radio-frequency identification. He received the IEEE Signal Processing Magazine Best Paper Award in 2007 and the EURASIP Technical Achievement Award in 2012. He is a Fellow of the IEEE.

REFERENCES

[1] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proc., F Radar Signal Process.*, vol. 140, no. 2, pp. 107–113, 1993.

[2] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," in *Handbook of Nonlinear Filtering*, E. D. Crisan and B. Rozovsky, Eds. Oxford, London: Oxford Univ. Press, 2011, pp. 656–704.

[3] A. Doucet, S. J. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. Comput.*, vol. 10, no. 3, pp. 197–208, 2000.

[4] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[5] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguez, "Particle filtering," *IEEE Signal Processing Mag.*, vol. 20, no. 5, pp. 19–38, 2003.

[6] O. Cappé, S. J. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *IEEE Proc.*, vol. 95, no. 5, pp. 899–924, 2007.

[7] G. Kitagawa, "Monte Carlo filter and smoother and non-Gaussian nonlinear state space models," *J. Comput. Graph. Stat.*, vol. 5, no. 1, pp. 1–25, 1996.

[8] J. Carpenter, P. Clifford, and P. Fearnhead, "An improved particle filter for nonlinear problems," *IEE Proc., Radar Sonar Navigat.*, vol. 146, no. 1, pp. 2–7, 1999.

[9] E. R. Beadle and P. M. Djurić, "A fast-weighted Bayesian bootstrap filter for nonlinear model state estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 1, pp. 338–343, 1997.

[10] J. Liu and R. Chen, "Sequential Monte-Carlo methods for dynamic systems," *J. Amer. Statist. Assoc.*, vol. 93, no. 443, pp. 1032–1044, 1998.

[11] J. S. Liu, R. Chen, and T. Logvinenko, "A theoretical framework for sequential importance sampling and resampling," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds. New York: Springer, 2001, pp. 225–246.

[12] M. Bolić, P. M. Djurić, and S. Hong, "Resampling algorithms for particle filters: A computational complexity perspective," *EURASIP J. Appl. Signal Process.*, vol. 2004, no. 15, pp. 2267–2277, 2004.

[13] R. Douc and O. Cappé, "Comparison of resampling schemes for particle filtering," in *Proc. 4th Int. Symp. Image and Signal Processing and Analysis*, 2005, pp. 64–69.

[14] J. D. Hol, T. B. Schön, and F. Gustafsson, "On resampling algorithms for particle filters," in *Proc. IEEE Nonlinear Statistical Signal Processing Workshop*, 2006, pp. 79–82.

[15] D. B. Rubin, "A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The SIR algorithm," *J. Amer. Statist. Assoc.*, vol. 82, no. 398, pp. 543–546, 1987.

[16] T. Li, T. P. Sattar, and S. Sun, "Deterministic resampling: Unbiased sampling to avoid sample impoverishment in particle filters," *Signal Process.*, vol. 92, no. 7, pp. 1637–1645, 2012.

[17] J. S. Liu, R. Chen, and W. H. Wong, "Rejection control and sequential importance sampling," *J. Amer. Statist. Assoc.*, vol. 93, no. 443, pp. 1022–1031, 1998.

[18] M. K. Pitt and N. Shephard, "Filtering via simulation: Auxiliary particle filters," *J. Amer. Statist. Assoc.*, vol. 94, no. 446, pp. 590–591, 1999.

[19] M. Bolić, P. M. Djurić, S. Hong, "New resampling algorithms for particle filters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2003, vol. 2, pp. 589–592.

[20] A. Budhiraja, L. Chen, and C. Lee, "A survey of numerical methods for nonlinear filtering problems," *Physica D*, vol. 230, no. 1–2, pp. 27–36, 2007.

[21] D. Crisan and T. Lyons, "A particle approximation of the solution of the Kushner-Stratonovitch equation," *Prob. Theory Relat. Fields*, vol. 115, no. 4, pp. 549–578, 1999.

[22] T. Li, T. P. Sattar, and D. Tang, "A fast resampling scheme for particle filters," in *Proc. Constantinides International Workshop on Signal Processing*, 2013, pp. 1–4.

[23] T. Li, S. Sun, and J. Duan, "Monte Carlo localization for mobile robot using adaptive particle merging and splitting technique," in *Proc. IEEE Int. Conf. Information and Automation*, 2010, pp. 1913–1918.

[24] P. Fearnhead and P. Clifford, "On-line inference for hidden Markov models via particle filters," *J. R. Statist. Soc.: Ser. B*, vol. 65, no. 4, pp. 887–899, 2003.

[25] J. Zheng, B. Bai, and X. Wang, "Increased-diversity systematic resampling in particle filtering for BLAST," *J. Syst. Eng. Electron.*, vol. 20, no. 3, pp. 493–498, 2009.

[26] D. Fox, "Adapting the sample size in particle filters through KLD-sampling," *Int. J. Robot. Res.*, vol. 22, no. 12, pp. 985–1003, 2003.

[27] T. Li, S. Sun, and T. Sattar, "Adapting sample size in particle filters through KLD-resampling," *Electron. Lett.*, vol. 46, no. 2, pp. 740–742, 2013.

[28] W. R. Gilks and C. Berzuini, "Following a moving target-Monte Carlo inference for dynamic Bayesian models," *J. R. Statist. Soc. B*, vol. 63, no. 1, pp. 127–146, 2001.

[29] C. Musso, N. Oudjane, and F. Legland, "Improving regularized particle filters," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds. New York: Springer, 2001, pp. 247–272.

[30] O. Hlinka, F. Hlawatsch, and P. M. Djurić, "Distributed particle filtering in agent networks: A survey, classification, and comparison," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 61–81, 2013.

[31] S. Hong, S. Chin, and P. M. Djurić, "Design and implementation of flexible resampling mechanism for high-speed parallel particle filters," *J. VLSI Signal Process.*, vol. 44, no. 1–2, pp. 47–62, 2006.

[32] M. Bolić, P. M. Djurić, and S. Hong, "Resampling algorithms and architectures for distributed particle filters," *IEEE Trans. Signal Processing*, vol. 53, no. 7, pp. 2442–2450, 2005.

[33] I. Strid, "Computational methods for Bayesian inference in macroeconomic models," Ph.D. dissertation, Stockholm School of Economics, 2010.

[34] O. Rosén, A. Medvedev, and M. Ekman, "Speedup and tracking accuracy evaluation of parallel particle filter algorithms implemented on a multicore architecture," in *Proc. IEEE Int. Conf. Control Applications*, 2010, pp. 440–445.

[35] G. Hendeby, R. Karlsson, and F. Gustafsson, "Particle filtering: The need for speed," *EURASIP J. Adv. Signal Process.*, vol. 2010, pp. 22:1–22:9, June 2010.

[36] K. Hwang and W. Sung, "Load balanced resampling for real-time particle filtering on graphics processing units," *IEEE Trans. Signal Processing*, vol. 61, no. 2, pp. 411–419, 2013.

[37] G. Hendeby, J. D. Hol, R. Karlsson, and F. Gustafsson, "A graphics processing unit implementation of the particle filter," in *Proc. European Signal Processing Conf.*, 2007, pp. 1639–1643.

[38] L. Murray, "GPU acceleration of the particle filter: The Metropolis resampler," arXiv preprint arXiv:1202.6163.

[39] A. C. Sankaranarayanan, A Srivastava, and R Chellappa, "Algorithmic and architectural optimizations for computationally efficient particle filtering," *IEEE Trans. Image Processing*, vol. 17, no. 5, pp. 737–748, 2008.

[40] B. Balasingam, M. Bolić, P. M. Djurić, and J. Míguez, "Efficient distributed resampling for particle filters," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 2011, pp. 3772–3775.

[41] J. Míguez, M. F. Bugallo, and P. M. Djuric´, "A new class of particle filters for random dynamical systems with unknown statistics," *EURASIP J. Appl. Signal Process.*, vol. 15, pp. 2278–2294, Jan. 2004.

[42] T. S. John, A. Nallanathan, and M. A. Armand, "A non-resampling sequential Monte Carlo detector for coded OFDM systems based on periodic termination of differential phase trellis," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1846–1856, 2006.

[43] J. Kotecha, and P. M. Djurić, "Gaussian particle filtering," *IEEE Trans. Signal Processing*, vol. 51, no. 10, pp. 2592–2601, 2003.

[44] T. Chen, T. B. Schön, H. Ohlsson, and L. Ljung, "Decentralized particle filter with arbitrary state decomposition," *IEEE Trans. Signal Processing*, vol. 59, no. 2, pp. 465–478, 2011.

[45] J. Míguez, "Analysis of parallelizable resampling algorithms for particle filtering," *Signal Process.*, vol. 87, no. 12, pp. 3155–3174, 2007.

[46] A. Kong, J. S. Liu, and W. H. Wong, "Sequential imputations and Bayesian missing data problems," *J. Amer. Statist. Assoc.*, vol. 9, no. 425, pp. 278–288, 1994.

[47] N. Celik and Y.-J. Son, "State estimation of a shop floor using improved resampling rules for particle filtering," *Int. J. Prod. Econ.*, vol. 134, no. 1, pp. 224–237, 2011.

[48] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Particle filters for joint timing and carrier estimation: Improved resampling guidelines and weighted Bayesian Cramér-Rao bounds," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1407–1419, May 2012.

[49] R. Douc and E. Moulines, "Limit theorems for weighted samples with applications to sequential Monte Carlo," *Ann. Stat.*, vol. 36, no. 5, pp. 2344–2376, 2008.

[50] P. Fearnhead, "Computational methods for complex stochastic systems: A review of some alternatives to MCMC," *Stat. Comput.*, vol. 18, no. 2, pp. 151–171, 2008.

[SP]
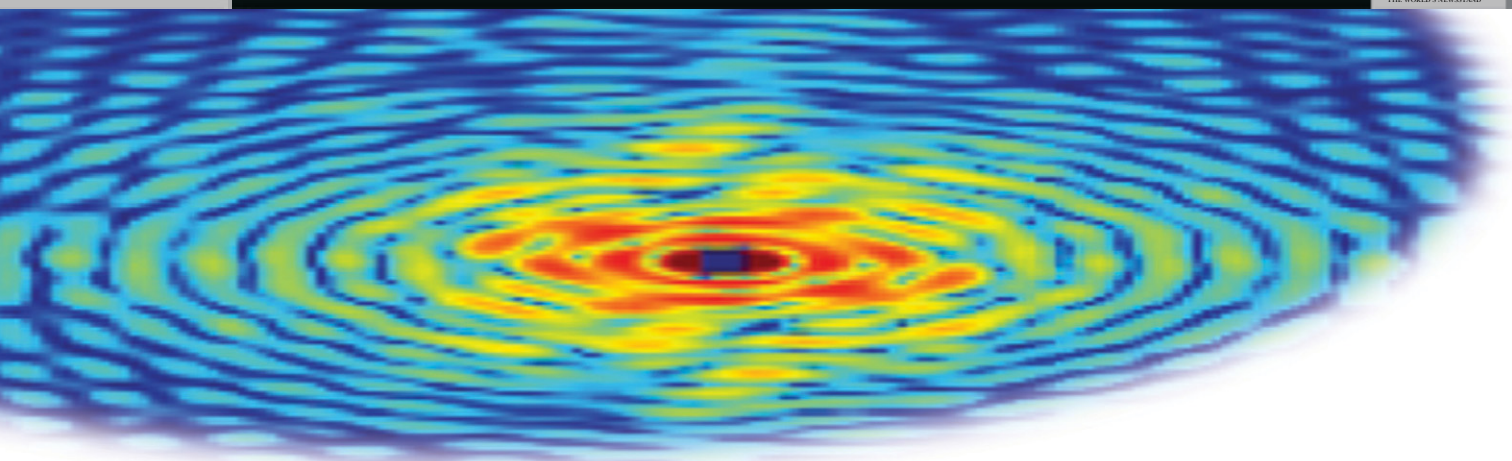
[Yoav Shechtman, Yonina C. Eldar, Oren Cohen, Henry N. Chapman, Jianwei Miao, and Mordechai Segev]

Phase Retrieval with Application to Optical Imaging

[A contemporary overview]

The problem of phase retrieval, i.e., the recovery of a function given the magnitude of its Fourier transform, arises in various fields of science and engineering, including electron microscopy, crystallography, astronomy, and optical imaging. Exploring phase retrieval in optical settings, specifically when the light originates from a laser, is natural since optical detection devices [e.g., charge-coupled device (CCD) cameras, photosensitive films, and the human eye] cannot measure the phase of a light wave. This is because, generally, optical measurement devices that rely on converting photons to electrons (current) do not allow for direct recording of the phase: the electromagnetic field oscillates at rates of $\sim 10^{15}$ Hz, which no electronic measurement device can follow. Indeed, optical measurement/detection systems measure the photon flux, which is proportional to the magnitude squared of the field, not the phase. Consequently, measuring the phase of optical waves (electromagnetic fields oscillating at 10^{15} Hz and higher) involves additional complexity, typically by requiring interference with another known field, in the process of holography.

Interestingly, electromagnetic fields do have some other features that make them amenable for algorithmic phase retrieval: their far field corresponds to the Fourier transform of their near field. More specifically, given a mask that superimposes an image on a quasi-monochromatic coherent field at some plane in space, the electromagnetic field distribution at a large enough distance from that plane is given

IEEE SIGNAL PROCESSING MAGAZINE [87] MAY 2015
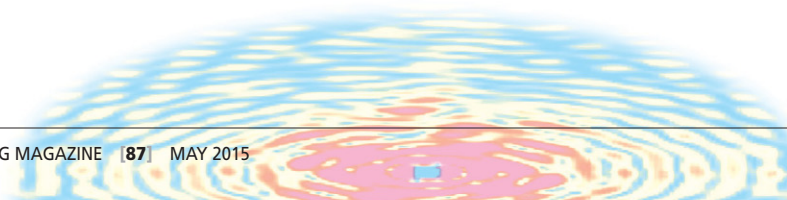
by the Fourier transform of the image multiplied by a known quadratic phase factor. Thus, measuring the far field, magnitude, and phase would facilitate recovery of the optical image (the wave field). However, as noted before, the optical phase cannot be directly measured by an electronic detector. Here algorithmic phase retrieval comes into play, offering a means for recovering the phase given the measurement of the magnitude of the optical far field and some prior knowledge.

This review article provides a contemporary overview of phase retrieval in optical imaging, linking the relevant optical physics to the signal processing methods and algorithms. Our goal is to describe the current state of the art in this area, identify challenges, and suggest future directions and areas where signal processing methods can have a large impact on optical imaging and on the world of imaging at large with applications in a variety of fields ranging from biology and chemistry to physics and engineering.

HISTORICAL BACKGROUND

Algorithmic phase retrieval offers an alternative means for recovering the phase of optical images without requiring sophisticated measuring setups as in holography. These approaches typically rely on some advanced information to facilitate recovery. In 1952, Sayre envisioned, in the context of crystallography, that the phase information of a scattered wave may be recovered if the intensity pattern at and between the Bragg peaks of the diffracted wave is finely measured [1]. In crystallography, the material structure under study is usually periodic (a crystal); hence, the far-field information contains strong peaks reflecting the Fourier transform of the usually periodic information. Measuring the fine features in the Fourier transform enabled the recovery of the phase in some simple cases. In 1978, 26 years later, Fienup developed algorithms for retrieving phases of two-dimensional (2-D) images from their Fourier modulus and constraints such as nonnegativity and a known support of the image [2] (see Figure 1).

In the early 1980s, the idea of phase retrieval created a flurry of follow-up work, partly because those times signified great hope for realizing an optical computer, of which phase retrieval was supposed to be a key ingredient. However, in the late 1980s and early 1990s, with the understanding that an optical computer is unrealistic, the interest in algorithmic phase retrieval diminished. Toward the end of the millennium, optical phase retrieval started to come back into contemporary optics research with the interest arising from a completely different direction: the community of researchers experimenting with X-ray imaging, where new X-ray sources (undulators and synchrotrons) were developed. The widespread interest in this field was mainly generated by the first experimental recording and reconstruction of a continuous diffraction pattern (Fourier magnitude squared) of a noncrystalline (nonperiodic) test object by Miao et al. in 1999 [3].

The reasons for the revival of optical phase retrieval in 1999 were actually quite subtle. One goal of optical imaging systems is to increase resolution, i.e., to image smaller and smaller features. However, as proved by Abbe's work in 1873, the highest attainable resolution in diffraction imaging (the so-called diffraction limit) is comparable to the wavelength of the light. For visible light, this diffraction limit corresponds to fractions of microns. Consequently, features on the molecular scale cannot be viewed with visible light in a microscope. One could argue then, why not simply use electromagnetic waves of a much shorter wavelength, say, in the hard X-ray regime, where the wavelength is comparable to atomic resolution? The reason is that lens-like devices and other optical components in this spectral region suffer from very large aberrations and are very difficult to make because refractive indices of materials in this wavelength regime are close to one. On the other hand, algorithmic phase retrieval is of course not limited by the quality of lenses; however, it requires very low noise detectors.

An additional problem is that as resolution is improved (i.e., as voxel elements in the recovered image are smaller in size), the number of photons per unit area must obviously increase to provide a reasonable signal-to-noise ratio (SNR). This means that the required exposure time to obtain a given signal level must increase as $(1/d)^4$, with d being the resolution length, assumed
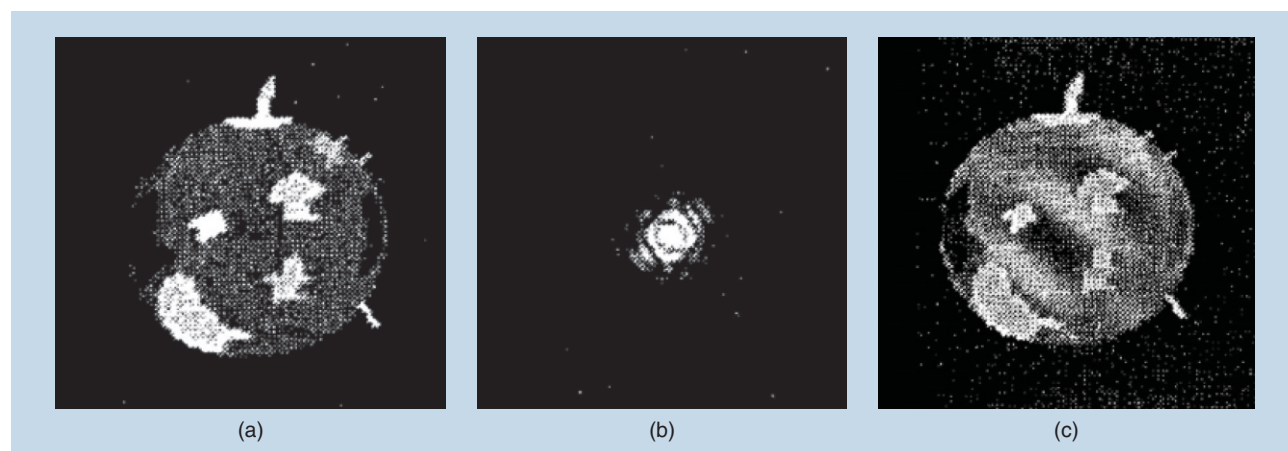
[FIG1] A numerical 2-D phase-retrieval example adapted from Fienup's 1978 paper [2]: (a) test object, (b) Fourier magnitude, and (c) reconstruction results [using hybrid input-output (HIO)—see Figure 3(b) for details]. (Images used with permission from [2].)

to be larger than atomic scales [4]. This, in turn, creates another problem: X-ray photons are highly energetic. The atomic cross section for photoabsorption is usually much higher than for elastic scattering, meaning that for every photon that contributes to the diffraction pattern (the measured Fourier magnitude), a considerable greater number of photons are absorbed by the sample. This energy dissipates in the sample first by photoionization and the breakage of chemical bonds, followed by a cascade of collisional ionization by free electrons and, at longer timescales, a destruction of the sample due to radiolysis, heating, or even ablation of the sample. Such radiation damage hinders the ability to recover the structure of molecules: the measured far-field intensity (Fourier magnitude) would reflect the structural damages, rather than providing information about the true molecular structure.

A solution to this problem was suggested by Solem and Chapline in the 1980s. They proposed to record images with pulses that are shorter than the timescale for the X-ray damage to occur. They predicted that picosecond pulses would be required to image at nanometer-length scales [5]. Toward the late 1980s, with the growing promise in constructing X-ray lasers that generate ultrashort pulses on the femtosecond scale, it was suggested that such pulses could even outrun damage processes at atomic length scales [6]. However, forming a direct image in this way would still require high-quality optical components (lenses and mirrors) in the X-ray regime, which do not currently exist. This is because creating lenses for the hard X-ray wavelength regime requires fabrication at picometer resolution, much smaller even than the Bohr radius of atoms. Likewise, while mirrors for X-rays do exist, their best resolution is on the scale of many nanometers, much larger than the features one would want to resolve in the imaging of molecules, for example.

The difficulties outlined earlier in direct X-ray imaging leave no choice but to use alternative methods to recover the structure of nanometric samples. Here is where phase retrieval can make its highest impact. Placing an area detector far enough from the sample to record the far-field diffraction intensity (which is approximately proportional to the squared magnitude of the Fourier transform of the image if the coherence length of the X-ray wave is larger than the sample size [7], [145]), together with appropriate constraints on the support of the sample, enable the recovery of the image at nanometric resolution. Indeed, the phase information has been shown numerically and experimentally to be retrieved in this fashion in various examples [2], [8]–[12].

The combination of X-ray diffraction, oversampling, and phase retrieval has launched the currently very active field called *coherent diffractive imaging* (*CDI*) [3]. In CDI, an object is illuminated with a coherent wave and the far-field diffraction intensity pattern (corresponding to the Fourier magnitude of the object) is measured. The problem then is to recover the object from the measured far-field intensity (see "Coherent Diffractive Imaging" and Figure S1). Since its first experimental demonstration, CDI has been applied to image a wide range of samples using synchrotron radiation [13]–[15], X-ray free-electron lasers (XFELs) [16], [17], high harmonic generation [18]–[20], soft X-ray laser [21], optical laser [22], and electrons [23], [24]. Recent reviews on the development and implementation

COHERENT DIFFRACTIVE IMAGING
In the basic CDI setup (forward scattering), an object is illuminated by a quasi-monochromatic coherent wave and the diffracted intensity is measured (Figure S1). When the object is small and the intensity is measured far away, the measured intensity is proportional to the magnitude of the Fourier transform of the wave at the object plane with appropriate spatial scaling.
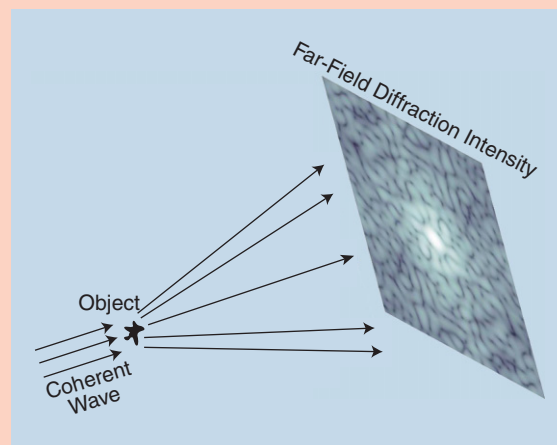
[FIGS1] A forward-scattering CDI setup: a coherent wave diffracts from an object (the sought information) and produces a far-field intensity pattern corresponding to the magnitude of the Fourier transform of the object.

In optics terms, when the Fresnel number is small ($N_F = (a^2/\lambda d) \ll 1$, where a is a radius confining the object in the object plane, d is the distance between the object and the measured intensity plane, and λ is the wavelength of the light), the relationship between the measured intensity I_{out} and the wave at the object plane E_{in} is given by [37]

$$I_{out}(x, y) \propto \left| \hat{E}_{in}\left(\frac{x}{\lambda d}, \frac{y}{\lambda d}\right) \right|^2$$

with $\hat{E}_{in} = \mathcal{F}\{E_{in}\}$ and $\mathcal{F}$ denoting the Fourier transform. Once the far-field intensity is measured, the goal is to recover E_{in} (which is equivalent to recovering the object) from I_{out}. This requires solving the phase-retrieval problem, which is attempted using an algorithm such as the ones described in this article.

of phase-retrieval algorithms for the specific application of CDI were written by Marchesini [9], Thibault and Elser [25], and Nugent [26]. Presently, one of the most challenging problems in CDI is three-dimensional (3-D) structural determination of large protein molecules [6]. There has been ongoing progress toward this goal over the past decade; see, e.g., [16], [17], [27], and [28].

Another research field where phase retrieval plays an important role is astronomy, where the objects are usually distant stars, which are optically incoherent sources. In such cases of incoherent waves, the phase is stochastic; hence, the optical signal is the

intensity of the light (amplitude of the complex field squared). This has important implications on algorithmic phase retrieval in terms of the assumptions that can be made on the signal (e.g., nonnegativity). One application of phase retrieval in astronomical measurements is for adaptive optics-based aberration correction, caused either by atmospheric turbulence or by imperfections in the optical imaging system [29]–[31]. Phase retrieval is also used in speckle interferometry [32], [33], a method to obtain information and later images [34], [35] beyond the diffraction limit of the (telescopic or alike) imaging system. As phase retrieval plays a major role in astronomy, there exist several detailed reviews from this perspective [31], [33], [36].

From a theoretical and algorithmic point of view, phase retrieval is a difficult problem, in many cases lacking a unique solution. Furthermore, even with the existence of a unique solution, there is not necessarily a guarantee that it can be found algorithmically. Nevertheless, as reasoned earlier, phase-retrieval algorithms and applications have benefited from a surge of research in recent years, in large part due to various new imaging techniques in optics. This trend has begun impacting the signal processing community as well—the past few years have witnessed growing interest within this community in developing new approaches to phase retrieval by using the tools of modern optimization theory [38], [39]. More recent work has begun exploring connections between phase retrieval and structure-based information processing [40]–[45]. For example, it has been shown that, by exploiting the sparsity of many optical images, one can develop powerful phase-retrieval methods that allow for increased resolution considerably beyond Abbe's diffraction limit, resolving features smaller than one-fifth of the wavelength [45]. The relationship between the fields of sparsity and optical imaging has led to an important generalization of the basic principles of sparsity-based reconstruction to nonlinear measurement systems [41], [44], [46]–[53]. Here too, optics played an important role in signal processing: since the phase-retrieval problem is inherently nonlinear (i.e., the signal is related to the measurements nonlinearly), employing sparsity-based concepts in phase retrieval required modifications to the linear sparsity-based algorithms known from the field of compressed sensing [54]. We believe that this field will grow steadily in the next few years, with rapid development of coherent X-ray sources worldwide [55], [56] and more researchers contributing to the theory, algorithms, and practice of nonlinear sparse recovery.

MATHEMATICAL FORMULATION

PROBLEM FORMULATION

Consider the discretized one-dimensional (1-D) real-space distribution function of an object: $\mathbf{x} \in \mathbb{C}^N$ (extension of the formulation to higher dimensions is straightforward). In CDI, for

> **THE RELATIONSHIP BETWEEN THE FIELDS OF SPARSITY AND OPTICAL IMAGING HAS LED TO AN IMPORTANT GENERALIZATION OF THE BASIC PRINCIPLES OF SPARSITY-BASED RECONSTRUCTION TO NONLINEAR MEASUREMENT SYSTEMS.**

example, this corresponds to the transmittance function of the object. The fact that $\mathbf{x}$ is generally complex corresponds physically to the fact that the electromagnetic field emanating from different points on the object has not only magnitude but also phase (as is always the case, for example, when 3-D objects are illuminated and light is reflected from points at different planes). The 1-D discrete Fourier transform (DFT) of $\mathbf{x}$ is given by

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi \frac{kn}{N}}, \qquad k = 0, 1, \ldots, N-1. \qquad (1)$$

The term *oversampled DFT* used in this article will refer to an M point DFT of $\mathbf{x} \in \mathbb{C}^N$ with $M > N$

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi \frac{kn}{M}}, \qquad k = 0, 1, \ldots, M-1. \qquad (2)$$

The recovery of $\mathbf{x}$ from measurement of $\mathbf{X}$ can be achieved by simply applying the inverse-DFT operator. Writing $X[k] = |X[k]| \cdot e^{j\phi[k]}$, the Fourier phase-retrieval problem is to recover $\mathbf{x}$ when only the magnitude of $\mathbf{X}$ is measured, i.e., to recover $x[n]$ given $|X[k]|$. Since the DFT operator is bijective, this is equivalent to recovering the phase of $X[k]$, i.e., $\phi[k]$—hence the term *phase retrieval*. Denote by $\hat{\mathbf{x}}$ the vector $\mathbf{x}$ after padding with $N-1$ zeros. The autocorrelation sequence of $\hat{\mathbf{x}}$ is then defined as

$$g[m] = \sum_{i=\max\{1,m+1\}}^{N} \hat{x}_i \overline{\hat{x}_{i-m}}, \qquad m = -(N-1), \ldots, N-1. \qquad (3)$$

It is well known that the DFT of $g[m]$, denoted by $G[k]$, satisfies $G[k] = |X[k]|^2$. Thus, the problem of recovering a signal from its Fourier magnitude is equivalent to recovering a signal from its autocorrelation sequence.

Continuous phase retrieval can be defined similarly to its discrete counterpart as the recovery of a 1-D signal $f(x)$ from its continuous Fourier magnitude

$$|F(\nu)| = \left| \int_{\mathbb{R}} f(x) \exp(-j2\pi\nu x) dx \right|.$$

Many objects of interest, such as electromagnetic fields, are usually described by continuous functions. However, since the data acquisition is digitized (by CCD cameras and alike), and the processing is done digitally, we mostly treat the discrete case here.

The Fourier phase-retrieval problem is as a special case of the more general phase-retrieval problem, where we are given measurements

$$y_k = |\langle \mathbf{a}_k, \mathbf{x} \rangle|^2, \qquad k = 1, \ldots, M, \qquad (4)$$

with $\mathbf{a}_k$ denoting the measurement vectors. In discrete 1-D Fourier phase retrieval, the measurement vectors correspond to $\mathbf{a}_k[n] = e^{-j2\pi(kn)/M}$. For mathematical analysis, it is often easier to
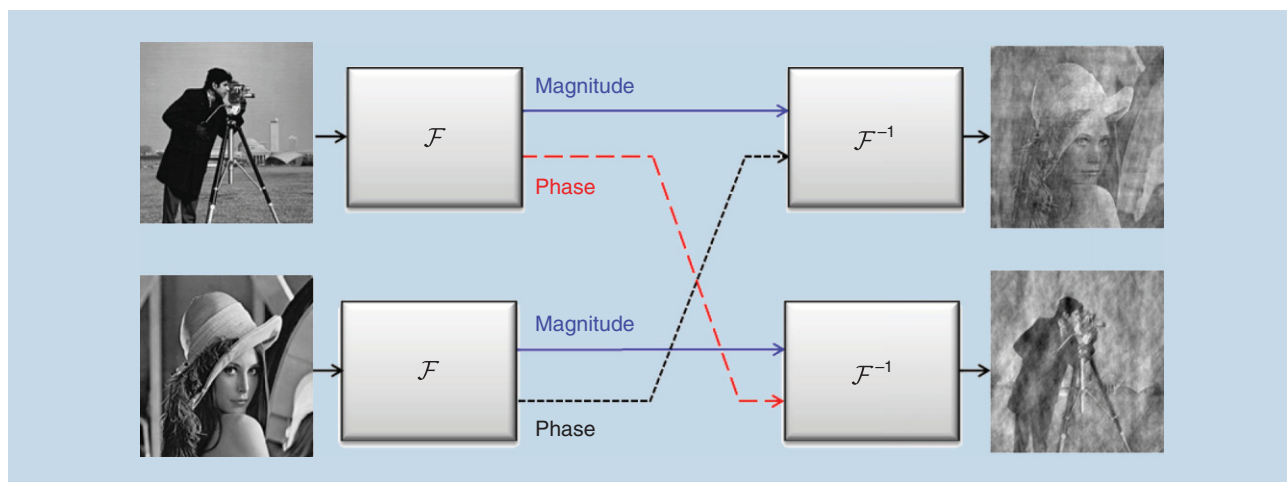
[FIG2] The importance of Fourier phase. Two images, a cameraman and Lenna, are Fourier transformed. After swapping their phases, they are inverse Fourier transformed. The result clearly demonstrates the importance of phase information for image recovery.

treat the case where the measurements are random (i.e., $\mathbf{a}_k$ are random vectors), as this allows uniqueness guarantees that are otherwise hard to obtain [38], [50], [57]–[59]. Nevertheless, more structured measurements have also been investigated [60].

Before proceeding to the mathematical methodology, it is important to highlight the significance of knowing the Fourier phase. In fact, it is well known that knowledge of the Fourier phase is crucial in recovering an object from its Fourier transform [61]. Many times the Fourier phase contains more information than the Fourier magnitude, as can be seen in the synthetic example shown in Figure 2. The figure shows the result of the following numerical experiment: two images (that of a cameraman and a woman named Lenna) are Fourier transformed. The phases of their transforms are swapped and, subsequently, they are inverse Fourier transformed. It is evident, for this quite arbitrary example, that the Fourier phase contains a significant amount of information about the images. In crystallography, this phenomenon is the source of genuine concern of phase bias of molecular models (such as those used in molecular replacement) in refined structures.

In the remainder of this section, we discuss uniqueness of the phase-retrieval problem, i.e., under what conditions the solution to the phase problem is unique. It is worth noting that, while the discussion of theoretical uniqueness guarantees is important and interesting, the lack of such guarantees does not prevent practical applications from producing excellent reconstruction results in many settings.

UNIQUENESS

FOURIER MEASUREMENTS

The recovery of a signal from its Fourier magnitude alone, in general, does not yield a unique solution. This section will review the main existing theoretical results regarding phase-retrieval uniqueness.

First, there are so-called trivial ambiguities that are always present. The following three transformations (or any combination of them) conserve Fourier magnitude:

1) global phase shift: $x[n] \Rightarrow x[n] \cdot e^{j\phi_0}$
2) conjugate inversion: $x[n] \Rightarrow \overline{x[-n]}$
3) spatial shift: $x[n] \Rightarrow x[n+n_0]$.

Second, there are nontrivial ambiguities, the situation of which varies for different problem-dimensions. In the 1-D setting, there is no uniqueness—i.e., there are multiple 1-D signals with the same Fourier magnitude. Even if the support of the signal is bounded within a known range, uniqueness does not exist [62]. Any pair of 1-D signals having the same autocorrelation function yields the same Fourier magnitude, as the two are connected by a Fourier transform. Consider, for example, the two vectors $\mathbf{u} = [1\,0-2\,0-2]^T$ and $\mathbf{v} = [(1-\sqrt{3})\,0\,1\,0\,(1+\sqrt{3})]^T$. Both of these vectors have the same support and yield the same autocorrelation function $g[m] = [-2, 0, 2, 0, 9, 0, 2, 0, -2]$. Therefore, they are indistinguishable by their Fourier magnitude, even though they are not trivially equivalent.

For higher dimensions (2-D and above), Bruck and Sodin [63], Hayes [64], and Bates [65] have shown that, with the exception of a set of signals of measure zero, a real $d \geq 2$ dimensional signal with support $\mathbf{N} = [N_1 \ldots N_d]$, i.e., $x[n_1, \ldots, n_d] = 0$ whenever $n_k < 0$ or $n_k \geq N_k$ for $k = 1, \ldots, d$ is uniquely specified by the magnitude of its continuous Fourier transform, up to the trivial ambiguities mentioned earlier. Furthermore, the magnitude of the oversampled $\mathbf{M}$ point DFT sequence of the signal, with $\mathbf{M} \geq 2\mathbf{N} - 1$ (where the inequality holds in every dimension), is sufficient to guarantee uniqueness. The problematic set of signals that are not uniquely defined by their Fourier magnitudes are those having a reducible Z transform: denoting the d-dimensional Z transform of x by $X(z_1, \ldots, z_d) = \sum_{n_1} \cdots \sum_{n_d} x[n_1, \ldots, n_d] z_1^{-n_1} \cdots z_d^{-n_d}$, $X(\mathbf{z})$ is said to be reducible if it can be written as $X(\mathbf{z}) = X_1(\mathbf{z}) X_2(\mathbf{z})$, where $X_1(\mathbf{z})$ and $X_2(\mathbf{z})$ are both polynomials in $\mathbf{z}$ with degree $p > 0$. It is important to note that, in practice, for typical images, a number of samples smaller than $2\mathbf{N} - 1$ is many times sufficient (even $\mathbf{M} \geq 2^{1/D}\mathbf{N}$ can work, where D is the dimension [66]); however, the exact guarantees relating the number of samples to the type of images remains an open question.

[TABLE 1] PHASE RETRIEVAL—UNIQUENESS.

FOURIER MEASUREMENT	1-D	NO UNIQUENESS [62]
	≥ 2D	UNIQUENESS FOR REAL NONREDUCIBLE SIGNALS. REQUIRES OVERSAMPLING BY ≈2 [64]
	k-SPARSE 1-D	UNIQUENESS FOR SIGNAL WITH COLLISION-FREE AUTOCORRELATION, (AND $K \neq 6$) [69] M FOURIER MAGNITUDE MEASUREMENTS ARE SUFFICIENT, FOR A PRIME $M \geq k^2 - k + 1$ [70]
GENERAL MEASUREMENTS	REAL SIGNAL R^N	SATISFYING THE COMPLEMENT PROPERTY IS NECESSARY AND SUFFICIENT. 2N-1 FULL-SPARK RANDOM MEASUREMENTS GUARANTEE UNIQUENESS WITH HIGH PROBABILITY [40]
	REAL SIGNAL R^N (NOISY)	N LOG (N) MEASUREMENTS (OR k LOG (N) MEASUREMENTS IN THE k-SPARSE CASE) ARE SUFFICIENT FOR STABLE UNIQUENESS [50]
	COMPLEX SIGNAL C^N	CONJECTURE: 4N−4 GENERIC MEASUREMENTS ARE SUFFICIENT [43]

Additional prior information about the signal, other than its support, can be incorporated and will naturally improve the conditioning of the problem. For example, knowledge of the Fourier phase sign (i.e., a single bit of phase information) has been shown [67] to yield uniqueness with some restrictions on the signal (specifically that the signal is real and its Z transform has no zeros on the unit circle). A different, popular, type of prior knowledge that has been used recently in various applications [54], [68] is that the signal $\mathbf{x} \in \mathbb{C}^N$ is sparse—i.e., contains only a small number k of nonzero elements, with $k \ll N$. The exact locations and values of the nonzero elements are not known a priori. In this case, it has been shown [69] that knowledge of the full autocorrelation sequence of a 1-D k-sparse real signal $\mathbf{x}$ is sufficient to uniquely define $\mathbf{x}$ as long as $k \neq 6$ and the autocorrelation sequence is collision free. A vector $\mathbf{x}$ is said to have a collision-free autocorrelation sequence if $x[i] - x[j] \neq x[k] - x[l]$, for all distinct $i, j, k, l \in \{1, \ldots N\}$ that are the locations of distinct nonzero values in $\mathbf{x}$. In addition, under these conditions, only M Fourier magnitude measurements are sufficient to uniquely define the autocorrelation sequence and, therefore, the signal $\mathbf{x}$, as long as M is prime and $M \geq k^2 - k + 1$ [70]. An interesting perspective relating phase retrieval to the Turnpike problem, for example, reconstructing a set of integers from their pairwise distances, is presented in [71]. Using this approach, the authors prove uniqueness with high probability for random signals having a nonperiodic support.

GENERAL MEASUREMENTS
Considering inner products with general non-Fourier (typically random) measurement vectors allows simpler derivation of theoretical guarantees. There have been several theoretical results relating the number and the nature of the measurements that are required for uniqueness, mostly dealing with random measurement vectors. The work of Balan [40] implies that, for real signals in $\mathbb{R}^N$, $2N-1$ random measurements are needed, provided that they are full spark, i.e., that every subset of N measurement vectors spans $\mathbb{R}^N$ [43]. This result was later extended to the complex case [43], where it is conjectured that $4N-4$ generic measurements, as defined in [43], are sufficient for bijectivity. In terms of stability, i.e., when the measurements are noisy, it is shown in [50] that on the order of $N \log(N)$ measurements [or $k \log(N)$ measurements in the

k-sparse case] are sufficient for stable uniqueness. Furthermore, minimizing the (nonconvex) least-squares objective: $\sum |y_i^2 - | \langle \mathbf{a}_i, \mathbf{x} \rangle |^2 |^p$, with $1 < p \leq 2$, yields the correct solution under these conditions [50]. For the noiseless case, any k-sparse vector in $\mathbb{R}^N$ has been shown to be uniquely determined by $4k - 1$ random Gaussian intensity measurements with high probability [70].

To study the injectivity of general (i.e., not necessarily random) measurements, the complement property was introduced in [40] for the real case. An extension was presented in [43] for the complex setting. A set of measurement vectors $\{\mathbf{a}_i\}_{i=1}^M$ with $a_i \in \mathbb{R}^N$ satisfies the complement property if for every $S \subseteq \{1, \ldots, M\}$, either $\{\mathbf{a}_i\}_{i \in S}$ or $\{\mathbf{a}_i\}_{i \in S^c}$ span $\mathbb{R}^N$. It has been shown in [40] that the mapping constructed by $y_i = |\langle \mathbf{a}_i, \mathbf{x} \rangle|$, $i = 1, \ldots, N$ is injective if and only if the measurement set satisfies the complement property. This poses a lower limit on the number of necessary measurements $M > 2N - 1$.

The results reviewed in this section are summarized in Table 1. In addition, there is a large amount of work on phase-retrieval uniqueness under different conditions, e.g., when the phase is known only approximately [72] or from redundant masked Fourier measurements [42], [73].

ALGORITHMS
Despite the uniqueness guarantees, no known general solution method exists to actually find the unknown signal from its Fourier magnitude given the other constraints. Over the years, several approaches have been suggested for solving the phase-retrieval problem, with the popular ones being alternating projection algorithms [2], [74], [75]. In addition, to help regularize the phase-retrieval problem, different imaging techniques were suggested that yield better behaved imaging models. For example, using exposures with different masks (e.g., the phase diversity method for aberration correction by adaptive optics [29], and also more recently [73]), or obtaining images at different propagation planes [31], [76], [77]. Another method to obtain additional information is scanning CDI (also termed *ptychography*) [78]–[80], which uses several different illumination patterns to obtain coherent diffraction images. Using such a modified imaging setup is then followed by applying an appropriate algorithm, performing the phase retrieval.

There are many existing approaches for phase retrieval. In this section, we focus on common general algorithms (see the

"General Algorithms" section) and sparsity-based methods, i.e., techniques exploiting prior knowledge in the form of signal sparsity (see the "Sparsity-Based Algorithms" section). We also discuss the transport-of-intensity equation (TIE) [81]–[83], which considers the recovery of an object's phase from several defocused intensity images.

GENERAL ALGORITHMS

The general phase-retrieval problem we wish to solve can be formulated as the following least squares problem or empirical risk minimization:

$$\min_{\mathbf{x}} \sum_{k=1}^{M} (y_k - |\langle \mathbf{a}_k, \mathbf{x} \rangle|^2)^2, \tag{5}$$

with $\mathbf{y}$ being the measurements and $\mathbf{a}_k$ being the measurement vectors defined in (4). In general, we can replace the square in the objective by any power p. Unfortunately, this is a nonconvex problem, and it is not clear how to find a global minimum even if one exists. In this section, we describe several approaches that have been suggested to deal with this problem and types of prior information that can be incorporated into these methods to increase the probability of convergence to the true solution.

ALTERNATING PROJECTIONS

The most popular class of phase-retrieval methods is based on alternate projections. These methods were pioneered by the work of Gerchberg and Saxton (GS) [74], dealing with the closely related problem of recovering a complex image from magnitude measurements at two different planes—the real (imaging) plane and Fourier (diffraction) plane. The original GS algorithm consists of iteratively imposing the real- and Fourier-plane constraints, such as the measured real-space magnitude $|x[n]|$, and Fourier magnitude $|X[k]|$, as illustrated in Figure 3(a). The GS iterations are described in Algorithm 1. The recovery error, defined as $E_i = \sum_k \left| |Z_i[k]| - |X[k]| \right|^2$, is easily

shown to be monotonically nonincreasing with i [75]. Despite this fact, recovery to the true solution is not guaranteed, as the algorithm can converge to a local minimum.

Algorithm 1: The GS algorithm.

Input: $|x[n]|, |X[k]|, \epsilon$
$|x[n]|$-Real-space magnitude
$|X[k]|$-Fourier magnitude
ϵ-Error threshold
Output: $z[n]$ - a vector that conforms with both magnitude constraints, i.e., : $|z[n]| = |x[n]|$, and $|Z[k]| = |X[k]|$, where $Z[k]$ is the DFT of $z[n]$
Initialization: Choose initial $z_0[n] = |x[n]| \exp(\phi[n])$ (e.g., with a random $\phi[n]$)
General Step: $(i = 1, 2, \ldots)$:
  1) Fourier transform $z_i[n]$ to obtain $Z_i[k]$
  2) Keep current Fourier phase, but impose Fourier magnitude constraint: $Z_i'[k] = |X[k]| \cdot Z_i[k]/|Z_i[k]|$.
  3) Inverse Fourier transform $Z_i'[k]$ to obtain $z_i'[n]$
  4) Keep current real-space phase, but impose real-space magnitude constraint: $z_{i+1}[n] = |x[n]| \cdot z_i'[n]/|z_i'[n]|$
  5) Go to 1
Until $E_i = \sum_k \left| |Z_i[k]| - |X[k]| \right|^2 \le \epsilon$

Extending the GS projection ideas further, in 1978 Fienup [2] suggested a modified version, in which the real-space magnitude constraints are replaced by other types of constraints, in addition to consistency with the measured Fourier magnitude. The real-space constraints may be, e.g., nonnegativity, a known signal support, i.e., $x[i] = 0$ for all $i > N_0$, where N_0 is known (or approximately known), or both. The basic framework of the Fienup methods is similar to GS—in fact, the first three steps are identical. Step 4, however, replaces imposing the real-space magnitude constraint by applying a correction to the real-space
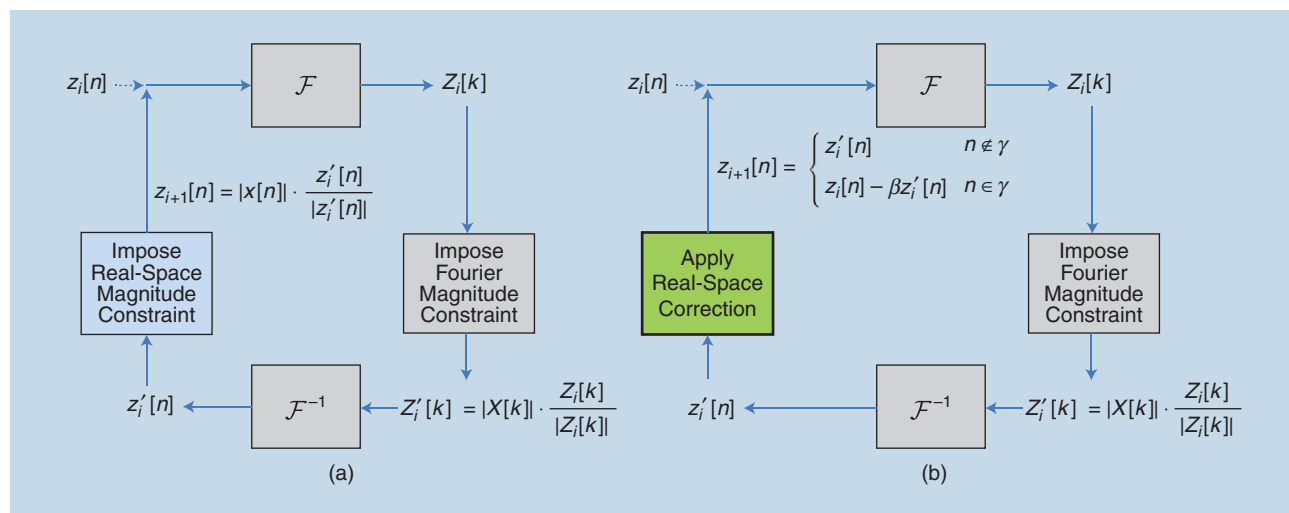
[FIG3] The block diagrams of (a) the GS algorithm and (b) the Fienup HIO algorithm. The algorithms differ in their fourth (colored) step.

estimate. Some possible variants to this step were also suggested [75]. Here, we describe the one most commonly used, referred to as the *hybrid input-output* (*HIO*) *method*, which consists of the following correction step:

4) Obtain $z_{i+1}[n]$ by applying a correction to the real-space image estimate:

$$z_{i+1}[n] = \begin{cases} z_i'[n], & n \notin \gamma \\ z_i[n] - \beta z_i'[n], & n \in \gamma, \end{cases} \tag{6}$$

with β being a small parameter and γ being the set of indices for which $z_i'[n]$ violates the real-space constraints.

The real-space constraint violation may be a support violation (e.g., a signal is nonzero where it should be zero) or a nonnegativity violation.

The Fienup algorithm is represented schematically in Figure 3(b). There is no proof that the HIO algorithm converges. It is also known to be sensitive to the accuracy of the prior information (e.g., the real-space support needs to be tightly known, especially in the complex signal case [84]). Nonetheless, in practice, the simple HIO-based techniques are commonly used in optical phase-retrieval applications such as CDI [85], [86]. Other variants of the correction step include the input–output method, and the output–output method [75], corresponding respectively to

$$z_{i+1}[n] = \begin{cases} z_i[n], & n \notin \gamma \\ z_i[n] - \beta z_i'[n], & n \in \gamma, \end{cases}$$
$$z_{i+1}[n] = \begin{cases} z_i'[n], & n \notin \gamma \\ z_i'[n] - \beta z_i'[n], & n \in \gamma. \end{cases} \tag{7}$$

An important feature of the HIO algorithm is its empirical ability to avoid local minima and converge to a global minimum for noise-free oversampled diffraction patterns. However, when there is high noise present in the diffraction intensity, HIO suffers from several limitations. First, the algorithm sometimes becomes stagnant and fails to converge to a global minimum. Second, a support has to be predefined. Third, the image oscillates as a function of the iteration. Over the years, various algorithms have been developed to overcome these limitations, including the combination of HIO and the error-reduction (ER) algorithm [75], difference map [8], hybrid projection reflection [10], guided HIO (GHIO) [87], relaxed averaged alternating reflectors (RAAR) [11], noise robust (NR)-HIO [88], and oversampling smoothness (OSS) [12].

As an example, the recently proposed OSS algorithm exhibits improved performance over HIO and its variants in many settings. OSS is based on Fienup iterations with an added smoothing Gaussian filter applied to the off-support region in the real-space object in each iteration. The fourth step in HIO is replaced by

$$z''[n] = \begin{cases} z_i'[n], & n \notin \gamma \\ z_i[n] - \beta z_i'[n], & n \in \gamma, \end{cases}$$
$$z_{i+1}[n] = \begin{cases} z_i''[n], & n \notin \gamma \\ \mathcal{F}\{Z_i''[k] W[k]\}, & n \in \gamma, \end{cases}$$

where $W[k]$ is a Gaussian function with its variance decreasing with iterations. A quantitative comparison for a specific example between OSS and HIO can be found in the section "Quantitative Comparison of Alternating-Projection Algorithms." For a comparison and numerical investigation of several alternate projection algorithms, see, e.g., [9] and [12].

The performance of Fienup methods is dependent on the initial points. Therefore, it is possible and recommended to try several initializations. In [58], the authors consider a clever method for initial point selection and show that for the random Gaussian measurement case, the resulting iterations yield a solution arbitrarily close to the true vector.

Analyses of iterative phase-retrieval algorithms from a convex optimization perspective can be found in [10] and [89]–[93]. In [91], the authors study the ER algorithm by viewing it as an iterated projections algorithm onto nonconvex sets. In [10] and [92], it is shown that the HIO method can be interpreted within different optimization frameworks depending on the constraints enforced. For example, given a support constraint, HIO coincides with the Douglas–Rachford algorithm for $\beta = 1$ [94], [95]. In [10], it is shown that under the same constraint, in the more general case of $\beta \neq 1$, HIO can be formulated in terms of projections and reflections. This representation, however, no longer holds when nonnegativity restrictions are added.

SEMIDEFINITE PROGRAMMING ALGORITHMS

An alternative recently developed to solve the phase-retrieval problem is based on semidefinite relaxation [39], [46], [57], [96]. The method relies on the observation that (4) describes a set of quadratic equations that can be rewritten as linear equations in a higher dimension. Specifically, define the $N \times N$ matrix $\mathbf{X} = \mathbf{xx}^*$. The measurements (4) are then linear in $\mathbf{X}$

$$y_k = |\langle \mathbf{a}_k, \mathbf{x} \rangle|^2 = \mathbf{x}^* \mathbf{a}_k \mathbf{a}_k^* \mathbf{x} = \mathbf{x}^* \mathbf{A}_k \mathbf{x} = \mathrm{Tr}(\mathbf{A}_k \mathbf{X}), \tag{8}$$

where $\mathbf{A}_k = \mathbf{a}_k \mathbf{a}_k^*$. Our problem is then to find a matrix $\mathbf{X} = \mathbf{xx}^*$ that satisfies (8). The constraint $\mathbf{X} = \mathbf{xx}^*$ is equivalent to the requirement that $\mathbf{X}$ has rank 1, and is positive semidefinite, which we denote by $\mathbf{X} \succeq 0$. Therefore, finding a vector $\mathbf{x}$ satisfying (4) can be formulated as

$$\begin{aligned} &\text{find} \quad \mathbf{X} \\ &\text{s.t.} \quad y_k = \mathrm{Tr}(\mathbf{A}_k \mathbf{X}), \quad k = 1, \ldots, M, \\ &\qquad \mathbf{X} \succeq 0, \\ &\qquad \mathrm{rank}(\mathbf{X}) = 1. \end{aligned} \tag{9}$$

Equation (9) is equivalent to the following rank minimization problem:

$$\begin{aligned} &\min \quad \mathrm{rank}(\mathbf{X}) \\ &\text{s.t.} \quad y_k = \mathrm{Tr}(\mathbf{A}_k \mathbf{X}), \quad k = 1, \ldots, M, \\ &\qquad \mathbf{X} \succeq 0. \end{aligned} \tag{10}$$

Unfortunately, rank minimization is a hard combinatorial problem. However, since the constraints in (10) are convex (in fact linear), one might try to relax the minimum rank objective, for

example, by replacing it with minimization of $\mathrm{Tr}(X)$. This approach is referred to as *PhaseLift* [39]. Alternatively, one may use the log-det reweighted rank minimization heuristic suggested in [97], which is the approach followed in [38] and [46]. In [38], it is shown that PhaseLift yields the true vector x with large probability when the measurements are random Gaussian and $M \sim O(N \log N)$.

An interesting approach is taken in [57], where x is separated into an amplitude component and a phase component, and only the phase is optimized. This approach yields several variations of existing methods, notably PhaseCut [57], which is a relaxation of the MaxCut algorithm [98] obtained by dropping the rank constraint.

The semidefinite programming (SDP) approach requires matrix lifting, i.e., replacing the sought vector with a higher-dimensional matrix, followed by solving a high-dimensional problem. It is, therefore, in principle, more computationally demanding than the alternating projection approaches, or greedy methods, which will be discussed in the section "Greedy Methods with Sparsity Prior." In addition, in general, there is no guarantee that the rank minimization process will yield a rank-1 matrix or that the true solution will be found even if there is a unique solution.

TRANSPORT OF INTENSITY

The TIE approach is a method that solves the known propagation equation of the electromagnetic field to recover the phase at some plane z_0, from several intensity measurements in the vicinity of that plane. Specifically, in the case of light propagation under the paraxial approximation (i.e., only small angles from the optical axis are considered, implying that the light field varies slowly on the scale of the optical wavelength), the TIE is

$$\frac{2\pi}{\lambda}\frac{\partial I}{\partial z} = -\nabla I \cdot \nabla \phi - I\Delta\phi, \qquad (11)$$

where $I(x, y, z_0)$ is the intensity distribution in plane z_0, λ is the wavelength of a monochromatic field, $\nabla = (\partial_x, \partial_y)$ is the transverse gradient, $\Delta = \nabla^2 = \partial_x^2 + \partial_y^2$ is the 2-D Laplacian, and $\phi(x, y, z_0)$ is the phase to be recovered. Recovering ϕ amounts to solving the partial differential equation (11). This can be achieved by first numerically estimating the derivative on the left-hand side of (11) using the measured intensity at two (or more) planes, e.g., $I(z_0)$ and $I(z_0 + dz)$, for a small dz. Then, after plugging in $I(z_0)$ into the right-hand side of (11), a variety of methods can be applied to solve for ϕ using appropriate

---

SPARSE LINEAR PROBLEMS

Finding sparse solutions to sets of equations is a topic that has drawn much attention in recent years [54], [68], [105], [106]. Consider the linear system

$$\mathbf{y} = \mathbf{Ax} \qquad (S1)$$

with **y** being a set of M linear measurements, **A** being an $M \times N$ measurement matrix, and **x** being the unknown length—N vector. When the system is underdetermined (i.e., $M < N$), there are infinitely many possible solutions **x**. A key result of the theory of sparse recovery is that adding the constraint that **x** is sparse, i.e., contains only a few nonzero entries guarantees a unique solution to (S1), under general conditions on **A**. One such condition is based on the coherence of **A** [107]

$$|\mathbf{x}|_0 \leq \frac{1}{2}\left(1 + \frac{1}{\mu}\right), \qquad (S2)$$

with $|\mathbf{x}|_0$ being the number of nonzero entries in **x**, and the coherence defined by

$$\mu = \max_{i,j} \frac{|<\mathbf{A}_i, \mathbf{A}_j>|}{\|\mathbf{A}_i\| \cdot \|\mathbf{A}_j\|}. \qquad (S3)$$

Here, we denote by $\mathbf{A}_i$ the ith column of **A**, and by $\|\mathbf{A}_i\|$ its Euclidean norm.

Under (S2), one can find the unique solution to (S1) by solving

$$\min_{\mathbf{x}}\|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{Ax}. \qquad (S4)$$

Unfortunately, (S4) is an NP-hard combinatorial problem. However, many methods have been developed to approximately solve (S4). One class of such methods consists of greedy algorithms such as orthogonal matching pursuit [108]. Another popular method is based on convex relaxation of the l_0 norm to an l_1 norm [109], which yields the convex problem

$$\min_{\mathbf{x}}\|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{Ax}. \qquad (S5)$$

In fact, under the condition (S2), it has been shown [107] that the solution to (S5) is equal to that of (S4).

Another important criterion to evaluate the recovery ability in sparse linear problems of the form (S1) is the restricted isometry property (RIP) [110] of **A**. For an $M \times N$ matrix **A** (with $M < N$), define the restricted isometry constant δ_k as the smallest value such that for every submatrix $\mathbf{A}_k$ composed of k columns of **A**

$$(1-\delta_k)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}_k\mathbf{x}\|_2^2 \leq (1+\delta_k)\|\mathbf{x}\|_2^2, \quad \forall \mathbf{x} \in \mathbb{R}^k. \qquad (S6)$$

The RIP is therefore a measure of whether **A** preserves the energy of any k-sparse signal—which is the case if δ_k is small. In the context of sparse recovery, it is used to prove uniqueness and noise-robustness results. For example, if **A** is such that $\delta_{2k} < \sqrt{2} - 1$, then solving (S5) will yield the unique sparse solution to (S1). In practice, it is combinatorially difficult to calculate the RIP of a given matrix. However, certain random matrices can be shown to have good RIP with high probability. For example, an $M \times N$ independent and identically distributed Gaussian matrix obeys the k-RIP with high probability, for $M \sim O(k \log(N/k))$ [105]. This is one of the reasons that random matrices are favorable for sparse sensing.

---

boundary conditions and further assumptions (a common one is that I is constant in x, y, so that $\nabla I = 0$ inside some boundary) [81]–[83], [99], [100].

The TIE approach requires acquisition of several images at different (and close) planes. It is relatively simple to implement when applicable and can produce phase measurements when the coherence of the light is not sufficient for interferometric measurements [101]. However, the necessity of multiple closely spaced imaging planes can naturally pose a limitation on possible applications, such as applications requiring a fast acquisition time or a high SNR. This is because multiple imaging planes require the use of beamsplitters, which leads to signal loss. Some tradeoffs between different parameters in the TIE approach, e.g., the amount of defocus (dz) versus recovery accuracy, are discussed in [102].

SPARSITY-BASED ALGORITHMS

A specific kind of prior knowledge that can be incorporated into the phase-retrieval problem to help regularize it is the fact that the sought real-space object is sparse in some known representation (see "Sparse Linear Problems"). This means that the object $\mathbf{x}$ can be written as

$$\mathbf{x} = \mathbf{\Psi}\boldsymbol{\alpha} \tag{12}$$

with $\mathbf{\Psi}$ being a representation matrix (the sparsity basis), and $\boldsymbol{\alpha}$ being a sparse vector, i.e., a vector containing a small number of nonzero coefficients. The simplest example is when the object is composed of a small number of point sources, in which case $\mathbf{\Psi}$ is the identity matrix. Equipped with such prior knowledge, one can hope to improve the performance of phase-retrieval algorithms by limiting the search for the true vector to the set of sparse vectors. There are several different ways that sparsity can be incorporated, which are described in this section.

ALTERNATING PROJECTIONS WITH SPARSITY PRIOR

The Fienup algorithm described in the section "Alternating Projections" allows, in principle, for the incorporation of various types of general knowledge about the object, including sparsity [41], [103]. Sparsity was shown to be a useful prior in phase-retrieval algorithms already in 2004 [104] in the iterative charge-flipping algorithm, although it was not exploited directly (the electron density in [104] is assumed to have extended regions of zeros). More explicitly, the method in [103], for example, is based on the Fienup iterations, with the first three steps remaining unchanged.
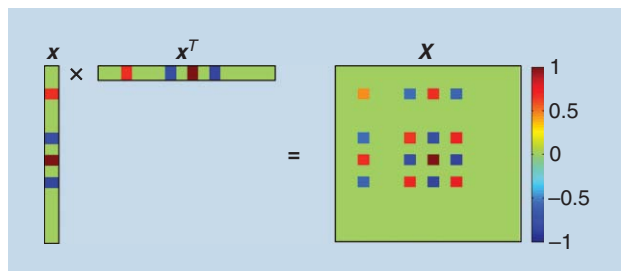
[FIG4] The sparse vector outer product yields a sparse matrix.

Step 4 is replaced by projection and thresholding. Assuming an invertible $\mathbf{\Psi}$ and a k-sparse vector $\boldsymbol{\alpha}$ such that $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\alpha}$:

4) Obtain $z_{i+1}[n]$ by projecting $z'_i[n]$ onto $\mathbf{\Psi}$, thresholding, and projecting back.
   - Calculate $\boldsymbol{\alpha}_i = \mathbf{\Psi}^{-1}\mathbf{z}'_i$.
   - Keep only the k largest elements of $|\boldsymbol{\alpha}_i|$, setting the rest to zero.
   - Set $\mathbf{z}_{i+1} = \mathbf{\Psi}\boldsymbol{\alpha}_i$.

Similar to the GS iterations, the error here can be shown to be nonincreasing so that convergence to a local minimum is guaranteed [103].

Note, that while this method is suggested in [103] for an orthonormal basis $\mathbf{\Psi}$, it can be easily modified to accommodate a noninvertible $\mathbf{\Psi}$. This can be done by replacing the first two parts with finding a sparse solution $\mathbf{a}_i$ to $\mathbf{z}_i = \mathbf{\Psi}\boldsymbol{\alpha}_i$, using any sparse solution approach [54].

SDP-BASED METHODS WITH SPARSITY PRIOR

SDP methods can also be modified to account for prior knowledge of signal sparsity. The incorporation of sparsity may be performed in several different ways. The first work to suggest sparsity-based SDP phase retrieval came from the domain of optics and dealt with partially spatially incoherent illumination [46]. This work actually considered a theoretical problem of greater complexity, combining phase retrieval with subwavelength imaging. Experimental results on subwavelength CDI can be found in [45], where the sought signal is an optical image with subwavelength features, and the measured data correspond to the Fourier magnitude sampled by a camera at the focal plane of a microscope lens.

The method suggested in [46], dubbed *quadratic compressed sensing* (QCS), is based on adding sparsity constraints to the rank minimization problem (10). When $\mathbf{x}$ is sparse, the result of the outer product $\mathbf{X} = \mathbf{x}\mathbf{x}^*$ is a sparse matrix as well, as shown in Figure 4. Therefore, one strategy might be to minimize the l_1 norm of the matrix $\mathbf{X}$. Alternatively, it is possible to further exploit the structure of $\mathbf{X}$ by noticing that the number of rows in $\mathbf{X}$ with a nonzero norm is equal to the number of nonzero values in $\mathbf{x}$. This means that the sparsity of $\mathbf{x}$ also implies a small number of nonzero rows in $\mathbf{X}$. Consider the vector $\mathbf{p}$ containing the l_2 norm of the rows of $\mathbf{X}$, i.e., $p_j = \left(\sum_k |X_{jk}|^2\right)^{(1/2)}$ (note that the l_2 norm can be replaced by any other norm). Since $\mathbf{p}$ should be sparse, one might try to impose a low l_1 norm on $\mathbf{p}$, in the spirit of l_1 minimization for the sparse linear problem. This yields the constraint $\|\mathbf{p}\|_1 = \sum_j |p_j| = \sum_j \left(\sum_k |X_{jk}|^2\right)^{(1/2)} \leq \eta$, corresponding exactly to a low mixed l_{1-2} norm constraint on $\mathbf{X}$ [111]. The problem to solve, as cast in [46], is therefore

$$
\begin{aligned}
\min \quad & \text{rank}(\mathbf{X}) \\
\text{s.t.} \quad & |\text{Tr}(\mathbf{A}_k\mathbf{X}) - y_k| \leq \epsilon, \quad k = 1, \ldots, M, \\
& \mathbf{X} \succeq 0, \\
& \sum_j \left(\sum_k |X_{jk}|^2\right)^{\frac{1}{2}} \leq \eta,
\end{aligned} \tag{13}
$$

where ϵ is a noise parameter and η is a sparsity parameter, enforcing row sparsity of $\mathbf{X}$.

Since finding a rank-1 matrix $\mathbf{X}$ satisfying the constraints is NP hard, the solution to (13) is approximated in [46] using the iterative log-det heuristic proposed in [97], with an additional thresholding step added at each iteration, to further induce signal sparsity. Once a low-rank matrix $\hat{\mathbf{X}}$ that is consistent with the measurements and the sparse prior is found, the sought vector $\mathbf{x}$ is estimated by taking the best rank-1 approximation of $\hat{\mathbf{X}}$ using the singular value decomposition: Decomposing $\hat{\mathbf{X}}$ into $\hat{\mathbf{X}} = \mathbf{U}\mathbf{S}\mathbf{U}^*$, the rank-1 approximation of $\hat{\mathbf{X}}$ is taken as $\hat{\mathbf{X}}_1 = S_{11}\mathbf{U}_1\mathbf{U}_1^*$, where S_{11} represents the largest singular value, and $\mathbf{U}_1$ is the corresponding column of $\mathbf{U}$.

Similar ideas that add sparse priors to SDP methods have been later suggested in [47], [57], and [112]. In [47], the rank minimization objective is relaxed to a convex trace minimization, with an additional l_1 regularization term to induce sparsity. This formulation yields

$$\min \quad \mathrm{Tr}(\mathbf{X}) + \lambda\|\mathbf{X}\|_1$$
$$\text{s.t.} \quad |\mathrm{Tr}(\mathbf{A}_k\mathbf{X}) - y_k| \le \epsilon, \quad k = 1, \ldots, M,$$
$$\mathbf{X} \succeq 0. \tag{14}$$

The solution of (14) is shown [47] to be unique in the noiseless case ($\epsilon = 0$), under the following condition: $\|\bar{\mathbf{X}}\|_0 \le (1/2)(1 + (1/\mu))$, where $\bar{\mathbf{X}} = \bar{\mathbf{x}}\bar{\mathbf{x}}^*$, with $\bar{\mathbf{x}}$ being the true solution to (4). The mutual coherence μ is defined by $\mu = \max_{i,j}(|<\mathbf{B}_i, \mathbf{B}_j>|)/(\|B_i\|\|B_j\|)$, with $\mathbf{B}$ being the matrix satisfying $\mathbf{y} = \mathbf{B}\mathbf{X}^\mathbf{S}$, where $\mathbf{X}^\mathbf{S}$ is the vector obtained from stacking the columns of $\mathbf{X}$. The same work also relates other recovery guarantees to the RIP criterion.

In [59] it is shown that for $\mathbf{a}_i$ that are independent, zero-mean normal vectors, on the order of $k^2 \log n$ measurements are sufficient to recover a k-sparse input from measurements of the form (4), using SDP relaxation. In [112], an algorithm is suggested to solve the sparse 1-D Fourier phase-retrieval problem based on a two-step process, with each step cast separately as an SDP problem: first, the support of $\mathbf{x}$ is determined from its autocorrelation sequence, and then $\mathbf{x}$ is found, given the support. This approach is shown experimentally to recover k-sparse signals from $O(k^2)$ measurements.

GREEDY METHODS WITH SPARSITY PRIOR

Since matrix-lifting algorithms involve a dimension increase, they are not ideally suited for large vectors, where the computational cost can become significant. In addition, they are generally not guaranteed to converge to a correct solution. An alternative is to use sparsity-based greedy methods [48], [51], [113]. One approach that is both fast and accurate is greedy sparse phase retrieval (GESPAR) [51]. GESPAR attempts to solve the least squares sparse quadratic problem (5). That is, it seeks a k-sparse vector $\mathbf{x}$ consistent with the quadratic measurements $\mathbf{y}$. It is a fast, local search method, based on iteratively updating the signal support, seeking a vector that corresponds to the measurements under the current support constraint. A local search method is repeatedly invoked, beginning with an initial random support set. Then, at each iteration, a swap is performed between a support and an off-support index. Only two elements are changed in the swap (one in the support and one in the off-support), following the so-called two-opt

method [114]. Given the support of the signal, the phase-retrieval problem is then treated as a nonconvex optimization problem, approximated using the damped Gauss Newton method [115]. See Algorithm 2 for a general description.

Algorithm 2: GESPAR—Main steps.

Input: $\mathbf{A}_i, y_i, \tau, \text{ITER}$.
$\mathbf{A}_i \in \mathbb{R}^{N \times N}, i = 1, 2, \ldots, M$ - symmetric matrices.
$y_i \in \mathbb{R}, i = 1, 2, \ldots, M$.
τ-threshold parameter.
ITER - Maximum allowed total number of swaps.
Output: $\mathbf{x}$-an optimal (or suboptimal) solution of (5).

Initialization: Set $T = 0$, $j = 0$.
   1) Generate a random index set S_0 ($|S_0| = s$)
   2) Invoke the damped Gauss–Newton method with support S_0 and obtain an output $\mathbf{z}_0$. Set $\mathbf{x}_0 = \mathbf{U}_{S_0}\mathbf{z}_0$, where $\mathbf{U}_{S_0} \in \mathbb{R}^{N \times s}$ is the matrix consisting of the columns of the identity matrix $\mathbf{I}_N$ corresponding to the index set S_0
General Step: ($j = 1, 2, \ldots$):
   3) Update support: Let p be the index from S_{j-1} corresponding to the component of $\mathbf{x}_{j-1}$ with the smallest absolute value. Let q be the index from S_{j-1}^c corresponding to the component of $\nabla f(\mathbf{x}_{j-1})$ with the highest absolute value, where $\nabla f(\mathbf{x})$ is the gradient of the least-squares objective function from (5), i.e., $\nabla f(\mathbf{x}) = 4\sum_i(\mathbf{x}^*\mathbf{A}_i\mathbf{x} - y_i)\mathbf{A}_i\mathbf{x}$. Increase T by 1, and make a swap between the indices p and q, i.e., set $\tilde{S}$ to be

$$\tilde{S} = (S_{j-1}\backslash\{p\}) \cup \{q\}.$$

   4) Minimize with given support: Invoke the damped Gauss–Newton method [115] with input $\tilde{S}$ and obtain an output $\tilde{\mathbf{z}}$. Set $\tilde{\mathbf{x}} = \mathbf{U}_S\tilde{\mathbf{z}}$, where $\mathbf{U}_S \in \mathbb{R}^{N \times s}$ is the matrix consisting of the columns of the identity matrix $\mathbf{I}_N$ corresponding to the index set S. If $f(\tilde{\mathbf{x}}) < f(\mathbf{x}_{j-1})$, then set $S_k = \tilde{S}$, $\mathbf{x}_k = \tilde{\mathbf{x}}$, and go to Step 3. If none of the swaps resulted with a better objective function value, go to Step 1.
Until $f(\mathbf{x}) < \tau$ or $T > \text{ITER}$.
The output is the solution $\mathbf{x}$ that yields the minimum value for the least-squares objective.

GESPAR has been shown to yield fast and accurate recovery results (see "Sparse Phase-Retrieval Algorithms—A Comparison" and Figure S2) and has been used in several phase-retrieval optics applications, including CDI of 1-D objects [116], efficient CDI of sparsely temporally varying objects [52], and phase retrieval via waveguide arrays [53]. A similar approach has been applied to treat the combined phase-retrieval and subwavelength imaging problem [45] (see the section "Subwavelength CDI Using Sparsity").

APPLICATIONS IN LENSLESS IMAGING

In this section, we present several CDI applications with connection to the phase-retrieval algorithms described previously. The

SPARSE PHASE-RETRIEVAL ALGORITHMS—A COMPARISON

We simulate sparse-Fienup [103] and GESPAR [51] for various values of $N \in [64, 2{,}048]$, and $M = 2N$. The recovery probability versus sparsity k for different vector lengths is shown in Figure S2(a) and (b). In both cases, the recovery probability increases with N, while GESPAR clearly outperforms the alternating iteration method.

We then simulate the recovery success rate of three sparsity-based phase-retrieval algorithms. We choose **x** as a random vector of length $N = 64$. The vector contains uniformly distributed values in the range $[-4, -3] \cup [3, 4]$ in k randomly chosen elements. The $M = 128$ point DFT of the signal is calculated, and its magnitude-square is taken as **y**, the vector of measurements. To recover the unknown vector **x**, three methods are used: a greedy method (GESPAR [51]), an SDP-based method [112, Algorithm 2], and an iterative Fienup algorithm with a sparsity constraint [103]. The sparse-Fienup algorithm is run using 100 random initial points, out of which the chosen solution is the one that best matches the measurements. $\hat{\mathbf{x}}$ is selected as the s-sparse output of the sparse-Fienup algorithm with the minimal cost $f(x) = \sum_{i=1}^{N} (|\mathbf{F}_i \mathbf{x}|^2 - y_i)^2$ out of the 100 runs. The probability of successful recovery is plotted in Figure 6(c) for different sparsity levels k. The success probability is defined as the ratio of correctly recovered signals **x** out of 100 simulations. In each simulation, both the support and the signal values are randomly selected. The three algorithms (GESPAR, SDP, and sparse-Fienup) are compared. The results clearly show that GESPAR outperforms the other methods in terms of probability of successful recovery—more than 90% successful recovery up to $k = 15$, versus $k = 8$ and $k = 7$ in the other two techniques. For more extensive comparisons, we refer the reader to [51].

A major advantage of greedy methods over other techniques (e.g., SDP based) is their low computational cost; GESPAR may be used to find a sparse solution to the 2-D Fourier phase retrieval—or phase retrieval of images. Figure S3 shows a recovery example of a sparse 195×195-pixel image comprised of $s = 15$ circles at random locations and random values on a grid containing 225 points, recovered from its 38,025 2-D Fourier magnitude measurements using GESPAR. The dictionary used in this example contains 225 elements consisting of nonoverlapping circles located on a 15×15-point Cartesian grid, each with a 13-pixel diameter. The solution took 80 s. Solving the same problem using the sparse-Fienup algorithm did not yield a successful reconstruction, and using the SDP method is not practical because of the large matrix size.
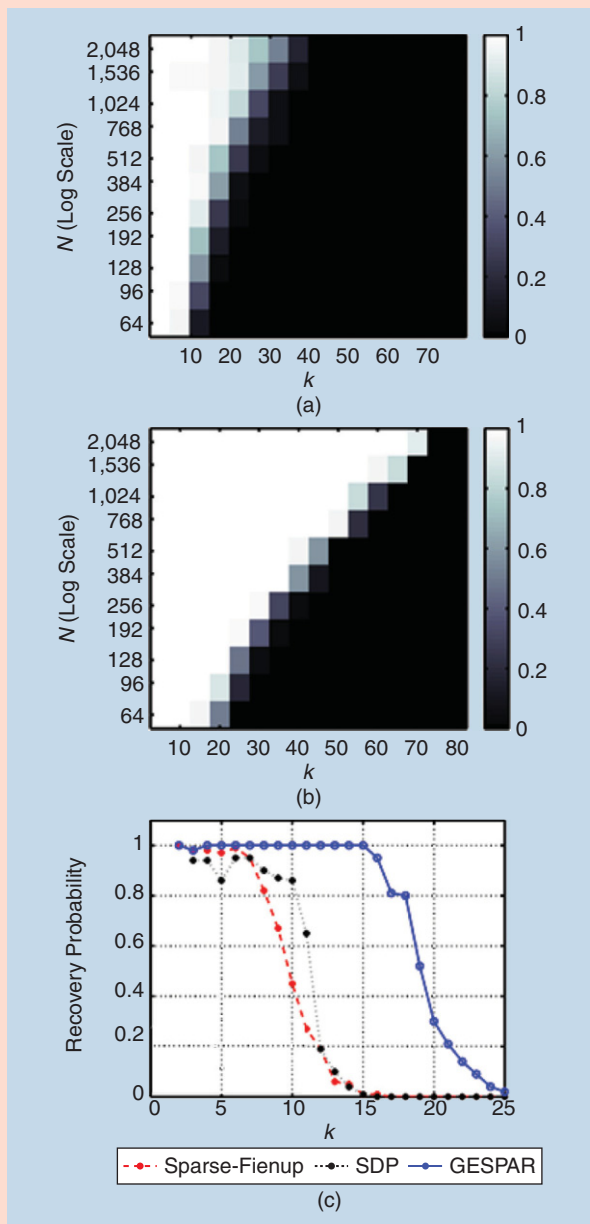
[FIGS2] A comparison of sparsity-based phase-retrieval algorithms. (a) The sparse-Fienup recovery probability versus sparsity k, for various signal length N, and with $M = 2N$. (b) GESPAR recovery probability versus sparsity k for various signal length N, and with $M = 2N$. (c) The recovery probability for three algorithms: sparse-Fienup, SDP, and GESPAR for $N = 64$ and $M = 128$ [51].
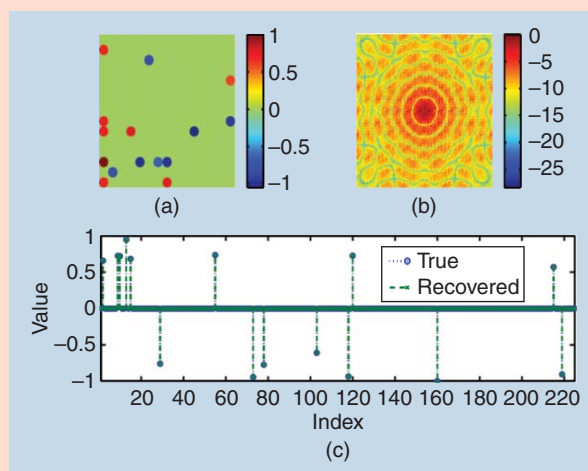
[FIGS3] A 2-D Fourier phase-retrieval example. (a) A true 195×195 sparse circle image ($s = 15$ circles). (b) The measured 2-D Fourier magnitude (38,025 measurements, log scale). (c) The true and recovered coefficient vectors corresponding to circle amplitudes at each of the 225 grid points [51].

concept of phase retrieval in optical imaging arises from the attempt to recover images from experimental measurements. To this end, it is essential to emphasize that, compared to numerical simulations or signal processing of digital data, phase retrieval of experimentally obtained patterns has several additional challenges. First, the far-field intensity distribution (Fourier magnitude) is corrupted by various types of noise, such as Poisson noise, detector readout noise, and unwanted parasitic scattering from the optics components in the system. Second, in single-shot experiments, the measured far-field intensity distribution is usually incomplete, including a missing center (i.e., the very low spatial frequency information cannot be directly recorded by a detector) [85]. Third, when the far-field intensity distribution is measured by a detector, each pixel integrates the total number of photons within the solid angle subtended by the pixel, which is not exactly equivalent to uniform sampling of the diffraction signal [117]. Additionally, many experiments are carried out using incoherent (but bright) sources. Spatial optical coherence [to distinguish from the term *coherence* in signal processing, as defined by (S3)] is achieved by propagating a long distance from the source, but often the experiment is constrained to be carried out with a partially incoherent beam [118].

All of these issues add complications to algorithmic phase retrieval. However, notwithstanding these challenges, successful

> **SINCE MATRIX-LIFTING ALGORITHMS INVOLVE A DIMENSION INCREASE, THEY ARE NOT IDEALLY SUITED FOR LARGE VECTORS, WHERE THE COMPUTATIONAL COST CAN BECOME SIGNIFICANT.**

phase retrieval of experimental data in optical imaging has been widely achieved [3], [13], [16], [17], [23], [28], [80], [119], [120]. Next we show several examples.

QUANTITATIVE COMPARISON OF ALTERNATING-PROJECTION ALGORITHMS

Quantitative comparisons between the OSS, HIO, ER-HIO, and NR-HIO algorithms have been performed using both simulated and experimental data [12]. Figure 5 shows a noise-free oversampled diffraction pattern (Fourier magnitude squared) calculated from a simulated biological vesicle [Figure 5(c)]. High Poisson noise was then added to the diffraction intensity [Figure 5(b)]. Figure 5(d)–(g) shows the final reconstructions by HIO, ER-HIO, NR-HIO, and OSS, respectively. Visually, OSS produced the most faithful reconstruction among the four algorithms [see the insets of Figure 5(d)–(g)]. The recovery error was quantified using consistency with the measurements

$$E = \sum_n |z_r[n] - z_m[n]| / \sum_n |z_m[n]|, \qquad (15)$$

where $z_r[n]$ is the final reconstruction and $z_m[n]$ is the model structure. The value for E of the HIO, ER-HIO, NR-HIO, and OSS reconstructions is 0.28, 0.24, 0.16, and 0.07, respectively.

Next, the four algorithms were compared using an experimental diffraction pattern measured from a *Schizosaccharomyces*
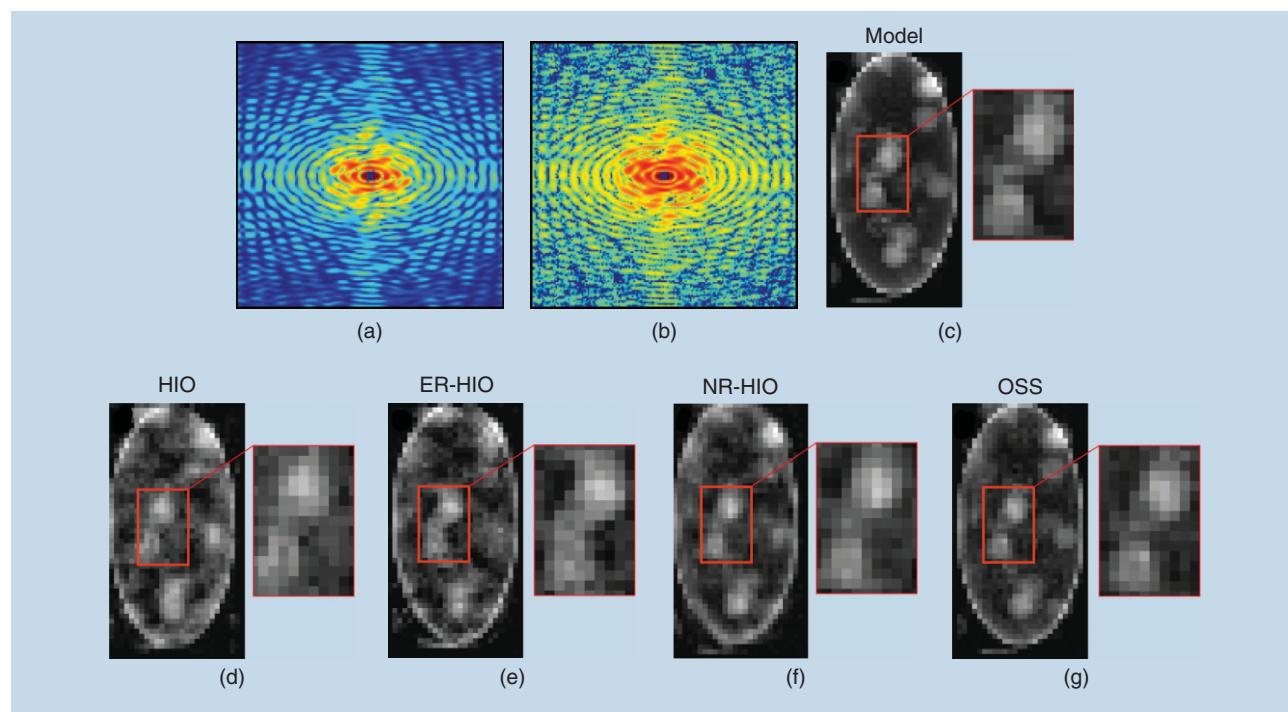
[FIG5] A quantitative comparison between the HIO, ER-HIO, NR-HIO, and OSS algorithms. (a) A noise-free oversampled diffraction pattern calculated from simulated biological vesicle. (b) The high Poisson noise added to the oversampled diffraction pattern. (c) The structure model of the biological vesicle and its fine features (inset). (b) The final reconstruction of the noisy diffraction pattern in (b) by (d) HIO, (e) ER-HIO, (f) NR-HIO, and (g) OSS [12].
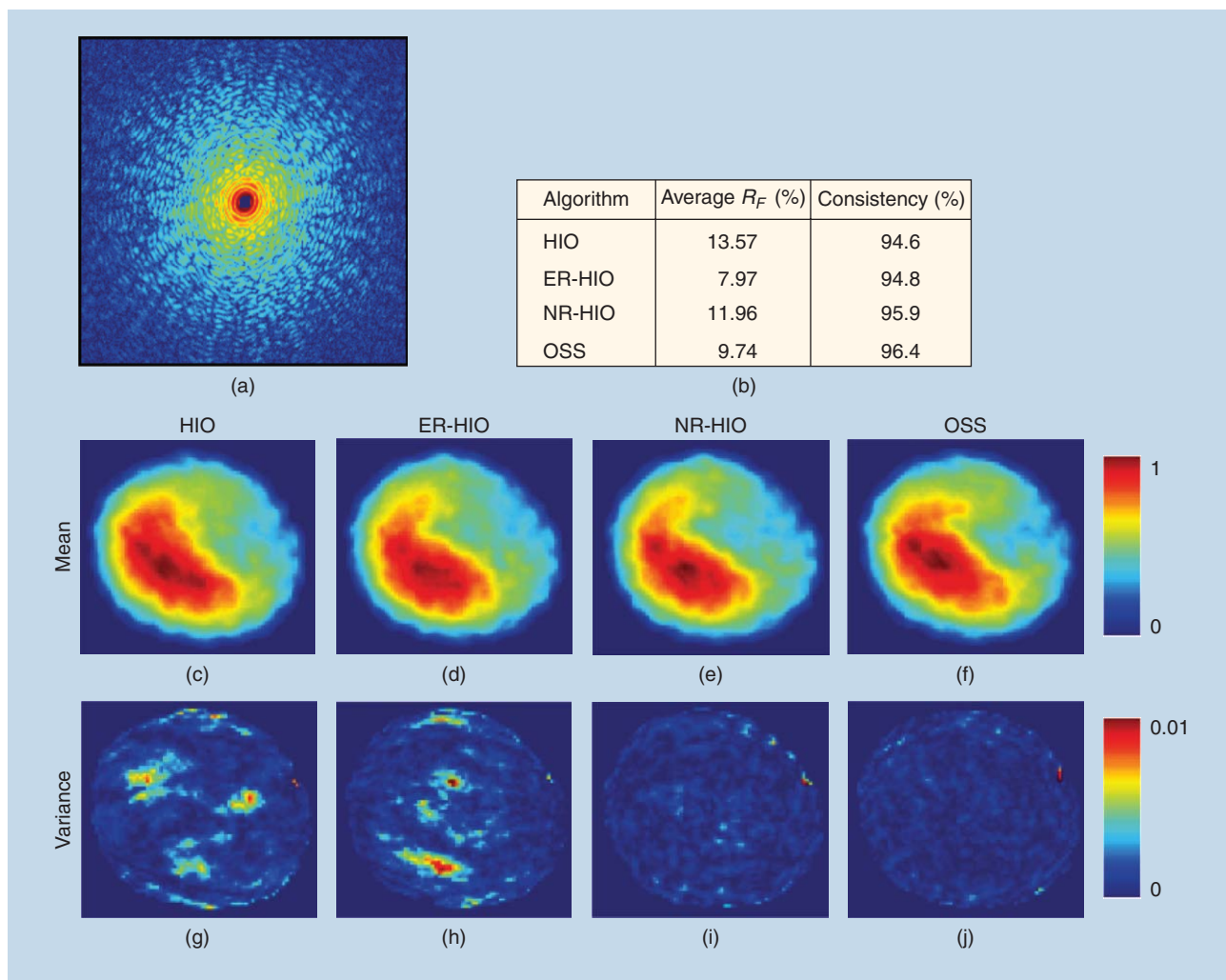
[FIG6] The phase retrieval of an experimental diffraction pattern from a biological sample. (a) An oversampled X-ray diffraction pattern measured from an *S. pombe* yeast spore cell. (b) The average R_F and the consistency of five independent trials of phase retrieval using four different algorithms. The average reconstruction of five independent trials using (c) HIO, (d) ER-HIO, (e) NR-HOP, and (f) OSS. The variance of five final images with (g) HIO, (h) ER-HIO, (i) NR-HOP, and (j) OSS [12].

pombe (*S. pombe*) yeast spore cell [12]. The experiment was conducted on an undulator beamline at a third-generation synchrotron radiation facility (Spring-8) in Japan. A coherent wave of 5 keV X-rays was incident on a fixed, unstained *S. pombe* yeast spore. An oversampled X-ray diffraction pattern was acquired by a CCD detector. Figure 6(a) shows the experimental diffraction pattern in which the centrosquare represents the missing low spatial resolution data [86]. By using a loose support, phase retrieval was performed on the measured data with the HIO, ER-HIO, NR-HIO, and OSS algorithms. For each algorithm, five independent trials were conducted, each consisting of 100 independent runs with different random initial phase sets. In each trial, the reconstruction with the smallest error metric R_F was chosen as a final image, where R_F is defined as

$$R_F = \sum_k \left| |Z_e[k]| - \zeta |Z_r[k]| \right| / \sum_k |Z_e[k]|. \tag{16}$$

Here, $|Z_e[k]|$ is the measured Fourier magnitude, $|Z_m[k]|$ is the recovered Fourier magnitude, and ζ is a scaling factor.

For each algorithm, the mean and average of the five final images were used to quantify the reconstruction. Figure 6(c)–(j) shows the average and variance of five final images obtained by HIO, ER-HIO [75], NR-HIO [88], and OSS [12], respectively. The average R_F and the consistency of five independent trials are shown in Figure 6(b). Both visual inspection and quantitative results indicate that OSS produced the most consistent reconstructions among all four algorithms.

XFEL CDI
The majority of imaging experiments at XFEL sources use the method of CDI. The lensless nature of CDI is actually an advantage when dealing with extremely intense and destructive pulses, where one can only carry out a single pulse measurement with each object (say, a molecule) before the object

disintegrates. In such cases, often one cannot use any optical components at all, because any component, e.g., a lens, would be severely damaged by the extremely high flux of X-ray photons, and the damaged components will distort the measured data. CDI solves these problems: it works without the need for optical components. In this vein, CDI also facilitates reliable imaging of moving objects. Indeed, in many experiments, the objects move (flow) across the X-ray beam, for example, when the X-ray laser beam hits a focused aerosol beam or nanoparticles in a liquid jet. In such an experiment, the particle density is usually adjusted so that the X-ray laser pulse is more likely to hit a single particle than several. A particle is hit by chance by a pulse, but this is not known until the diffraction pattern is read out from the detector, which is done on every pulse. The stream of data is then analyzed and sorted to give the single-particle hits, which contain the meaningful measured data, while all other data are ignored.

There are two generic classes of these "single particle" CDI experiments: imaging of reproducible particles and imaging of unique particles. The first category includes particles such as viruses. Assuming that these particles are not aligned in the same direction, the collected data represent diffraction patterns of a common object, but in random orientations. If the orientations

> SPATIAL OPTICAL COHERENCE IS ACHIEVED BY PROPAGATING A LONG DISTANCE FROM THE SOURCE, BUT OFTEN THE EXPERIMENT IS CONSTRAINED TO BE CARRIED OUT WITH A PARTIALLY INCOHERENT BEAM.

can be determined, then the full 3-D Fourier magnitude of the object can be determined, which in turn could be phased to give a 3-D image. A proof of concept of this experiment was carried out by Loh et al. [121].

An example of the second class of flash diffractive imaging is imaging airborne soot particles in flight in an aerosol beam [28]. Several diffraction patterns of soot particles and clusters of polystyrene spheres as test objects are shown in Figure 7 along with the 2-D reconstructions of the objects. The experiments were carried out at the Linac Coherent Light Source using the Center for Free-Electron Laser Science-Advanced Study Group multipurpose instrument [122] at the atomic, molecular, and optical science beam line [123]. Pulses of about 10^{12} photons of 1.0-nm wavelength were focused to an area of $10\ \mu m^2$. The X-ray detectors (pnCCD panels) were placed to give a maximum full-period resolution of 13 nm at their center edges.

In these experiments, the phase retrieval of the patterns was carried out using the RAAR [11] algorithm and shrinkwrap procedure [124], which determines and iteratively updates the support constraint used. The objects were such that it was possible to apply an additional constraint that the image is real valued. Strikingly, the X-ray coherent diffraction patterns have a very high contrast. The intensity minima are close to zero. This has an

[FIG7] The diffraction patterns from single X-ray FEL pulses from particles in flight and reconstructed images: (a)–(d) clusters of polystyrene spheres with radii of (a) and (b) 70 nm, and (c) and (d) 44 nm; (e) and (f) ellipsoidal nanoparticles; (g) a soot particle; and (h) a salt–soot mixture [28].
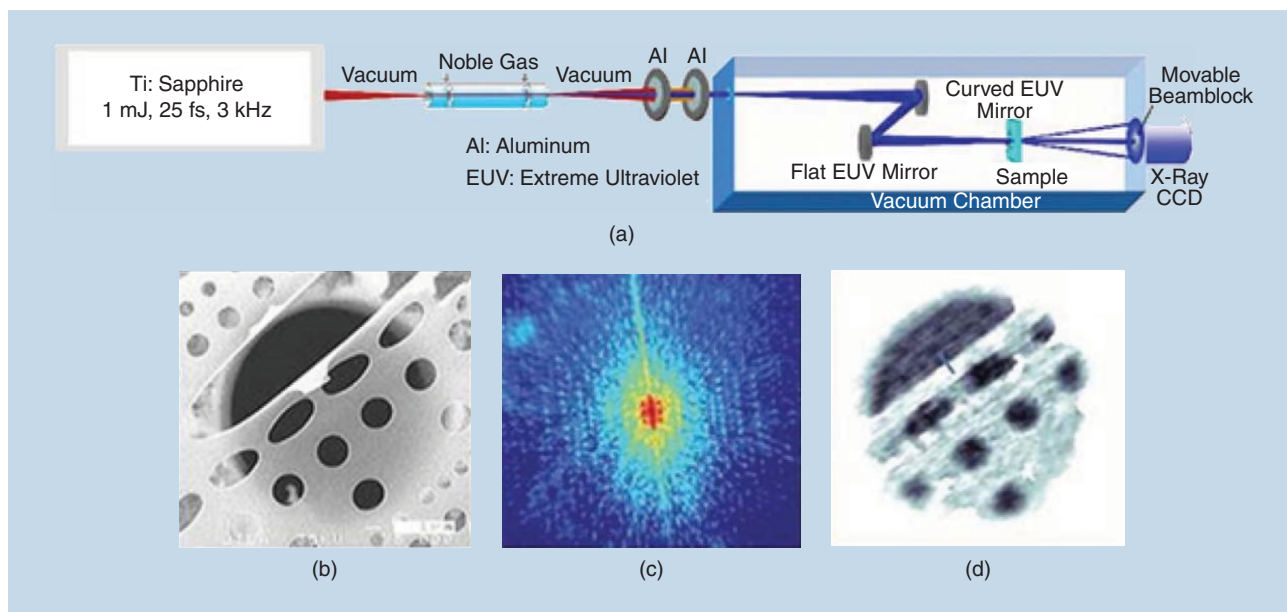
[FIG8] The first tabletop short-wavelength CDI. (a) The experimental setup. Coherent extreme UV radiation is generated through the process of high harmonic generation. A single harmonic order at wavelength 29 nm is selected and focused onto a sample by a pair of multilayer mirrors. The scattered light is detected by the X-ray CCD camera. (b) The original image, used to analyze the performance of the CDI process, obtained with a scanning electron microscope (SEM). The image shows a masked carbon film placed on a 15-μm diameter pinhole. (c) The recorded multiframe diffraction pattern [corresponding to Fourier magnitude squared of the object shown in (b)]. (d) CDI reconstruction using the GHIO algorithm with 214-nm resolution [18].

enormous effect on the ability to recover the phase of these patterns reliably. This reliability is quantified in the phase-retrieval transfer function (PRTF) [125], which compares the magnitude of the complex-valued average of patterns phased with different starting guesses to the square root of the measured diffraction pattern. If, at a particular pixel of the diffraction pattern, the phases are consistently reconstructed, then the sum over N patterns will give a magnitude N times higher than the measured magnitude, and so the PRTF will be unity. If the phases are random, then this sum will be close to zero. For patterns generated with XFELs, this function is often close to unity and is lower primarily in areas where the SNR is low.

Because the signal is limited, ultimately, so is the resolution; an estimate of the achieved resolution is given by the white dotted circle on each pattern in Figure 7. The reconstructed images are sums of ten independent reconstructions. These complex-valued sums have the property that their Fourier spectrum is effectively modulated by the PRTF and, hence, any artifact due to noise (or even due to forcibly truncating the data to a lower resolution) is unlikely to show up in the recovered image.

TABLETOP SHORT WAVELENGTH CDI
To date, most CDI experiments are carried out in third-generation synchrotron and XFELs. However, limited access and experimental time hinder the development and applications of CDI using these methods. Thus, over the past several years, CDI microscopes based on tabletop sources of coherent extreme ultraviolet and soft X-rays are also being developed [126]. Figure 8 shows the first tabletop CDI experiment with extreme UV wavelength.

Phase retrieval, i.e., obtaining Figure 8(d) from (c), is achieved using the GHIO algorithm [87]. In GHIO, the standard HIO is first run in parallel starting from several (16 in this case) random initial points, for a set number of iterations (2,000). This is generation zero of the algorithm. Then, the best output (in the sense of distance from the measurements) is selected to serve as the seed for the next generation. The inputs for the first generation are
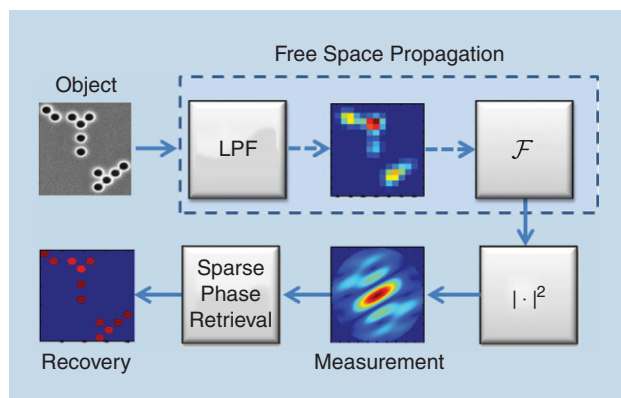
[FIG9] The sparsity-based subwavelength CDI. A 2-D object consisting of an arrangement of nanoholes (100 nm in diameter) is illuminated by a 532-nm laser, and the Fourier plane magnitude is measured. High spatial frequencies are lost during propagation, because the features (the circles as well as their separation) are smaller than $\sim\lambda/2$. Using an iterative greedy algorithm, and exploiting the prior knowledge that the object is sparse in a dictionary made of 100-nm circles, the phase is retrieved and the object is recovered from its low-pass-filtered Fourier magnitude [45].

obtained by multiplying the seed with each of the 16 images and taking the square root of the product. The same procedure is repeated for the next generations. The result in Figure 8(d) was obtained after the eighth generation, for which the 16 reconstructed images became consistent.

SUBWAVELENGTH CDI USING SPARSITY

Prior knowledge of object sparsity can help regularize the phase-retrieval problem as well as compensate for loss of other kinds of information. Here, we consider a problem in which the high spatial frequencies are lost. As described before, when an object is illuminated by coherent light of wavelength λ, the far-field intensity pattern is proportional to the magnitude of the object's Fourier transform. In addition, features in the object that are smaller than $\sim\lambda/2$ are smeared due to the diffraction limit. Consequently, the intensity measured in the far field corresponds to $y \propto |\mathbf{LFx}|^2$, where $\mathbf{L}$ represents a low-pass filter at cutoff frequency $\nu_c = 1/\lambda$, $\mathbf{F}$ represents the Fourier transform, and $|\cdot|^2$ stands for elementwise squared absolute value.

Figure 9 (adapted from [45]) shows the recovery of a sparse object containing subwavelength features (100-nm holes illuminated by a $\lambda = 532$-nm laser) from its experimentally measured low-pass-filtered Fourier magnitude. The prior knowledge used for recovery is that the object comprises a small number of 100-nm diameter circles on a grid, illuminated by a plane wave. The exact number, locations, and amplitudes of the circles are not known a priori. The recovery is performed using a greedy algorithm that iteratively updates the support of the object, finds a local minimum, and removes the weakest circle until convergence [45].

Another type of information loss in CDI, for which the prior knowledge of object sparsity can be helpful, is low SNR. In nondestructive X-ray CDI measurements, it is not uncommon for signal acquisition time to be on the order of tens of seconds [18], [20], [127] to achieve sufficiently high SNR. This poses a severe limitation on the temporal resolution attainable with such measurements, restricting the types of dynamical phenomena accessible by X-ray CDI. Exploiting sparsity in the change that an object undergoes between subsequent CDI frames has been recently suggested as a means to overcome high noise values and, consequently, significantly decrease acquisition time [52]. In other words, the fact that an object is sparsely varying can be used as prior information to effectively denoise sequential Fourier magnitude measurements. In [52], CDI of a sparsely varying object is formulated as a sparse quadratic optimization problem and solved using GESPAR [51]. Numerical simulations suggest a dramatic potential improvement in temporal resolution. In an example consisting of a 51×51-pixel object with five randomly varying pixels between frames, an improvement of two orders of magnitude in acquisition time is possible [52].

Finally, in [53], an experimental proof of concept is presented for an optical system in which discrete phase retrieval is performed using a small number of intensity measurements. The system considered is a model multiple-input, multiple-output communication system: an array of coupled optical waveguides in which a small (sparse) number of input waveguides is excited. As the light propagates through the array, the energy couples into neighboring waveguides until, ultimately, at the output plane, the energy is distributed among many of the waveguides. The purpose is to recover the complex input field, i.e., which waveguides were excited, and at what amplitude and phase, given output intensities of only a subset of the waveguides. This problem is formulated as a discrete phase-retrieval problem, and the loss of information, both of phase and of unmeasured waveguides, is compensated by a sparsity prior. The phase is then retrieved using GESPAR [51].

> KNOWLEDGE OF THE FOURIER PHASE SIGN HAS BEEN SHOWN TO YIELD UNIQUENESS WITH SOME RESTRICTIONS ON THE SIGNAL.

OTHER PHYSICAL SETTINGS, BOTTLENECKS, AND VISION

This review article is focused on the simplest physical setting for phase retrieval in optical imaging (Figure 2), CDI: an unknown 2-D optical image is recovered algorithmically from a single measurement of its far-field intensity pattern, given a known image support (or other prior information). In terms of signal processing, this problem corresponds to recovering a 2-D object from measurements of its Fourier magnitude. However, the issue of phase retrieval in optical imaging and, in a more general sense, in optics is far broader and includes other physical settings that naturally translate into signal processing problems that are different than the standard phase-retrieval formulation. This section provides a short overview of those physical settings, defines the various problems in terms of signal processing, and provides some key references. We conclude with a discussion on the main challenges and bottlenecks of phase retrieval in optical imaging, followed by an outlook for the upcoming years and long-term vision.

NON-FOURIER MEASUREMENTS

The simplest optical phase-retrieval problem assumes that the measured data corresponds to the Fourier magnitude. In optical settings, this means that the measurements are taken in the Fourier domain of the sought image, which physically means performing the measurements at a plane sufficiently far away from the image plane (the so-called far field or the Fraunhofer regime) or at the focal plane of an ideal lens [37]. In reality, however, the measurements can be taken at any plane between the image plane and the far field, which would yield intensity patterns that are very different than the Fourier magnitude of the signal. This of course implies that new (or revised) algorithms—beyond those described in previous sections—must be used, which naturally raises issues of conditions for uniqueness and convergence. At the same time, these measurements have some interesting advantages, which can be used wisely to improve the performance of phase retrieval. Let us begin by describing the relevant physical settings.

As stated earlier, the optical Fourier plane corresponds to a plane sufficiently far away from where the object (the sought signal) is

positioned. Far away here means asymptotically at infinite distance from the object plane or at the focal plane of a lens. However, the entire propagation-evolution of electromagnetic waves from any plane to any other plane is known: it is fully described by Maxwell's equations. As such, one can formulate the problem through a proper transfer function of the electromagnetic wave that is different than the Fourier transform.

In this context, the most well-studied case is the regime of Fresnel diffraction, where the transfer function is expressed in an integral form known as the *Fresnel integral* [37]. This regime occurs naturally at a range of distances away from the object plane, which naturally also includes the Fraunhofer regime where the transfer function reduces to a simple Fourier transform. Going beyond the Fresnel regime is also possible. This means that the (magnitude squared of the) electromagnetic wave will be measured at some arbitrary plane away from the object. A more general case arises by expressing the scalar transfer function of the light in a homogeneous medium, at any plane z as

$$T(k_x, k_y, z) = \exp\left[-iz\sqrt{k^2 - (k_x^2 + k_y^2)}\right]. \qquad (17)$$

Here, $k = \omega/c$, with ω being the frequency of the light, c being the speed of light in the medium, and k_x, k_y describe the transverse wavenumbers. The field at any arbitrary plane z, $E(x, y, z)$, is then given by inverse Fourier transforming the spectral function at that plane $F(k_x, k_y, z)$ [namely, the Fourier transform of $E(x, y, z)$ with respect to x, y], which is related to the spectrum at the initial plane by

$$F(k_x, k_y, z) = F(k_x, k_y, z = 0)\, T(k_x, k_y, z).$$

With the transfer function (17), one can now formulate a new phase-retrieval problem, where the measurements are conducted at some arbitrary plane z, giving $|E(x, y, z)|^2$, and the sought signal is $E(x, y, z = 0)$. This approach can be extended to include polarization effects, in which the transfer function is vectorial, thereby describing the propagation through Maxwell's equations with no approximation at all. The optical far field—where the measurement corresponds to the Fourier magnitude of the image at the initial plane (i.e., the measurement is proportional to $|F(k_x, k_y, z)|^2$)—is obtained for distances z larger than some minimum distance z_0 that depends on the spectral extent of $F(k_x, k_y, z = 0)$, and only within a region close enough to the z-axis in the measurement plane.

It is interesting to compare these more general phase-retrieval problems to the generic problem of recovering a signal from its Fourier magnitude. In terms of algorithmics, the generic problem is much simpler and was extensively studied throughout the years, whereas the general case is considerably more complex and was studied only sporadically. However, in terms of optics, the measurements in the general case can provide more information.

Namely, measurements of $|E(x, y, z)|^2$ can be taken at multiple planes (multiple values of z), and each measurement adds more information on the signal. In contrast, for the generic problem, once the measurements are taken in the optical far field, taking more measurements at further away distances does not add additional information because all of the far-field measurements correspond to the Fourier magnitude (to within some known scaling of coordinates in the measurement planes). As such, performing phase retrieval of optical images in the most general (non-Fourier) case can be beneficial as it leads to multiple measurements, possibly relaxing the conditions on oversampling and/or the advance knowledge on the support in the image plane.

> **THE MULTIFREQUENCY IDEA ALSO WORKS IN THE CONTINUOUS CASE OF BROAD BANDWIDTH PULSES CENTERED ON A SINGLE FREQUENCY.**

Historically, these ideas on non-Fourier measurements have been known to the optics community since the early days of optical phase retrieval [2]. They are currently being used in the context of improving the convergence of phase retrieval by taking non-Fourier measurements at several planes [14], [128]. Alternatively, one can take measurements at several different optical frequencies ω, which would be expressed as different values of $k = \omega/c$ in the general transfer function given before. In this multifrequency context, it is important that the frequencies are well separated, each having a narrow bandwidth, to conform the high degree of coherence required for CDI. These ideas are now being pursued by several groups [19], [118], [129]. Interestingly, the multifrequency idea also works in the continuous case of broad bandwidth pulses centered on a single frequency. In this case, the power spectrum of the pulse must be known in advance [118], [129], [130]. In a similar vein, recent work has demonstrated scanning CDI, where the beam is scanned through overlapping regions on the sample to allow imaging of extended objects, a method known as *ptychography* [80], [131]–[133].

More sophisticated physical settings also exist, where the medium within which the waves are propagating is not homogeneous in space. Famous examples are photonic crystals, wherein the refractive index varies periodically in space, in a known fashion, in one, two, or three dimensions. Obviously, in such settings, the transfer function for electromagnetic waves is fundamentally different from the transfer function in free space. The phase-retrieval problem in such systems, albeit less commonly known, is no less important. For example, photonic crystal fibers can in principle be used for imaging in endoscopy. The measurements in such systems correspond to the magnitude squared of the field at the measurement plane, which would be very different than the Fourier magnitude of the image. Still, once the transfer function is known, complicated as it may be, the phase-retrieval problem is well defined and can be solved with some modifications to the algorithms described earlier; see, e.g., the pioneering work on phase retrieval in a photonic crystal fiber [134], and very recently, work on sparsity-based phase retrieval and superresolution in optical waveguide arrays [53]. In addition to these, the concept of CDI has

also been extended to other schemes, such as Bragg CDI, suitable to periodic images to reconstruct the structure and strain of nanocrystals [135]–[138].

PHASE RETRIEVAL COMBINING HOLOGRAPHIC METHODS

As explained earlier, optical settings always suffer from the inability of photodetectors to directly measure the phase of an electromagnetic wave at frequencies of terahertz and higher. A partial solution for this problem is provided through holography, which was invented by Denis Gabor in 1948 [139], who was awarded the Nobel Prize in Physics in 1971. Holography involves interfering an electromagnetic field carrying some image, E_{image}, with another electromagnetic field of the same frequency and a known structure, denoted as E_{ref}. Typically, the so-called reference wave, E_{ref}, has a very simple structure, for example, approximately a plane wave (wave of constant amplitude and phase). The detection system records $|E_{\text{image}} + E_{\text{ref}}|^2$. Originally, such holographic recording was done on a photographic plate that was made from a photosensitive material whose transmission, being sensitive to the intensity of the light, became proportional to the recorded pattern $|E_{\text{image}} + E_{\text{ref}}|^2$. Such a photographic plate is called a *hologram*, wherein the information contained in the image wave E_{image} is embedded in transmission function of the hologram. To see the recording, the wave of the known pattern, E_{ref}, is generated (which is possible because its structure is simple and fully known) and made to illuminate the hologram. The magnitude of the wave transmitted through the illuminated hologram is therefore proportional to $|E_{\text{image}} + E_{\text{ref}}|^2 \cdot E_{\text{ref}}$. One of the terms is $|E_{\text{ref}}|^2 \cdot E_{\text{image}}$. Since $|E_{\text{ref}}|^2$ carries virtually no information, i.e., it is just a constant, this transmitted wave reconstructs the image times that constant. This is the principle of operation of holography. Over the years, it has been shown that it is almost always beneficial to record not the actual image but its Fourier spectrum; hence, the reconstructed information is the Fourier transform of the image, and the image itself is recovered either in the far field (as explained in the beginning of this article) or at the focal plane of a lens. This process is termed *Fourier holography* [140].

In the context of phase retrieval, holography is used to add information in the measurement scheme. Because in most cases the measurements used are Fourier magnitudes, which physically imply far-field measurements, the natural inclusion of holographic methods is through Fourier holography. For example, adding a tiny hole (a delta function) at a predetermined position in the sample, close to where the sought image resides, creates an additional wave in the far field with a tilted phase that arises from the displacement between the hole and the sought image. The far-field intensity, therefore, now corresponds to the absolute value squared of the sum of the Fourier transform of the sought image and the known wave. As such, it introduces additional prior knowledge that can be used for increased resolution of the algorithmic recovery or for

relaxing the constraints on the prior knowledge on the support. These ideas have been exploited successfully using X-rays and electrons by several groups [141]–[143].

CHALLENGES

The current challenges can be briefly defined as higher resolution, the ability to recover more complex objects, improved robustness to noise, and real-time operation. The very reason phase retrieval in optical imaging has recently become so important is owing to the vision to be able to one day directly image complex biological molecules, track their structural evolution as it evolves over time, and even view the dynamics of the electronic wave functions bonding atoms together. The reasoning is obvious: to understand biology at the molecular level and to decipher the secrets of how their atomic constituents bond together

> **THE CURRENT CHALLENGES CAN BE BRIEFLY DEFINED AS HIGHER RESOLUTION, THE ABILITY TO RECOVER MORE COMPLEX OBJECTS, IMPROVED ROBUSTNESS TO NOISE, AND REAL-TIME OPERATION.**

and how they interact with other molecules. The current state of the art is far from those goals: imaging resolution is not yet at the atomic (subnanometer) level, and—at nanometric resolution— imaging cannot handle objects that are bounded by a support that is extremely large compared to the resolution. In terms of being able to perform real-time experiments, state-of-the-art measurements have demonstrated extremely short optical pulses: tens of attoseconds (10^{-18} s, on the order of the passage of a photon through a distance comparable to the size of an atom). Pioneering experiments have even started to probe the dynamics of electrons in molecules and tunneling processes on these timescales. But, as of today, none of these ultrafast methods was applied to imaging of even a simple molecule, let alone complex biological structures.

Clearly, the underlying physics and engineering pose great challenges to meet these goals. Generating coherent radiation in the hard X-ray regime is still a major obstacle, often requiring very large enterprises such as the X-ray sources at the SLAC National Accelerator Laboratory. These facilities around the world are continuously improving their photon flux at shorter wavelengths, thereby constantly improving imaging resolution. The fundamental limits on the coherent X-ray flux possible with current methods (such as synchrotrons, XFELs [55], [56], and the process of high harmonics generation [144]) are not even known. But the steady improvement does give hope for imaging at the atomic level in the near future. Taking the CDI techniques to the regime of attosecond science is an important challenge. These pulses are extremely short, and, hence, their bandwidth is huge. Therefore, the coherent diffraction pattern is a superposition of their multispectral contents, which requires new algorithmic methods. As described earlier, these issues are currently being explored by several groups. But the problem is fundamentally more complicated because the process of scattering of light by molecules at these short wavelengths and ultrashort timescales is not like passing light through a mask on which an image is imprinted. Rather, many issues related to light-matter interactions under these conditions are yet to be understood (e.g., tunneling ionization of atoms by laser pulses).

Finally, the long-term vision must include imaging the dynamics within complex biological systems at the atomic level and in real time. But such systems are extremely complex to handle, in terms of details on many spatial and temporal scales simultaneously, in terms of the statistical nature and huge redundancy in the physical processes taking place within such complexes simultaneously, and even in terms of the quantum mechanics governing the dynamics at those scales. This is where the signal processing community can make a large impact—by devising new and original methods for recovering the information from experimental measurements. Clearly, the algorithms will have to be tailored to the specific physical settings.

AUTHORS

Yoav Shechtman (yoavsh@stanford.edu) received his B.Sc. degree in electrical engineering and physics in 2007 and his Ph.D. degree in physics in 2013 from the Technion–Israel Institute of Technology, Haifa, Israel. He is currently a postdoctoral scholar in the Department of Chemistry at Stanford University. His research interests include optical information processing, sparse information processing, single-molecule tracking, and fluorescence-based superresolution imaging. He is specifically interested in applying signal processing methods to practical biological/physical optical imaging problems. He has been awarded the Irwin and Joan Jacobs Excellence Scholarship for graduate studies and research, the Gutwirth Excellence Scholarship for graduate studies, and the Hershel Rich Innovation Award.

Yonina C. Eldar (yonina@ee.technion.ac.il) is a professor in the Department of Electrical Engineering at the Technion–Israel Institute of Technology, Haifa, Israel, and holds the Edwards Chair in Engineering. She has received numerous awards for excellence in research and teaching, including the Wolf Foundation Krill Prize for Excellence in Scientific Research, the Hershel Rich Innovation Award, the Michael Bruno Memorial Award from the Rothschild Foundation, the Weizmann Prize for Exact Sciences, the Muriel and David Jacknow Award for Excellence in Teaching, the IEEE Signal Processing Society Technical Achievement Award, and the IEEE/AESS Fred Nathanson Memorial Radar Award. She is the editor-in-chief of *Foundations and Trends in Signal Processing*. She is a Fellow of the IEEE. She is also a member of the Young Israel Academy of Science and the Israel Committee for Higher Education.

Oren Cohen (oren@tx.technion.ac.il) is an associate professor in the Department of Physics at the Technion–Israel Institute of Technology, Haifa, Israel. His research interests are mainly in extreme nonlinear optics, including high harmonic generation, laser filamentation, and spatiotemporal solitons, and also in structure-based superresolution in imaging, spectroscopy, and diagnostics of ultrashort laser pulses. He has received several awards including the Wolf Foundation Krill Prize for Excellence in Scientific Research and the Hershel Rich Innovation Award.

Henry N. Chapman (henry.chapman@desy.de) is the director of the Center for Free-Electron Laser Science at the Deutsches Elektronen-Synchrotron and the University of Hamburg in Germany. He earned his Ph.D. degree in X-ray optics at the University of Melbourne, Australia. He has developed methods in coherent X-ray imaging, which began at Stony Brook University in New York. At Lawrence Livermore National Laboratory in California, he led a team to demonstrate that an intense X-ray free-electron laser pulse could outrun radiation damage. He continued this work to the atomic scale with the method of serial femtosecond X-ray crystallography, which promises to overcome current bottlenecks in protein structure determination. His current research is focused on developing this method and extending it to the smallest possible crystals, i.e., single molecules.

Jianwei Miao (miao@physics.ucla.edu) is a professor at the University of California, Los Angeles. He received an M.S. degree from the Chinese Academy of Science in 1994 and a Ph.D. degree from Stony Brook University in 1999. He conducted a seminal experiment on extending X-ray crystallography to allow structural determination of noncrystalline specimens in 1999. This method, known as *coherent diffractive imaging*, has been broadly implemented using synchrotron radiation, X-ray free-electron lasers, high harmonic generation, and electrons. In 2012, he demonstrated electron tomography at 2.4-Angstrom resolution, the highest resolution ever achieved in any general three-dimensional (3-D) imaging method. More recently, he applied electron tomography to observe nearly all the atoms in a platinum nanoparticle, and for the first time imaged the 3-D core structure of edge and screw dislocations at atomic resolution.

Mordechai Segev (msegev@tx.technion.ac.il) received his B. Sc. and Ph.D. degrees from the Technion–Israel Institute of Technology, Haifa, Israel, in 1985 and 1990. After three years of postdoctoral research at the California Institute of Technology, he joined Princeton University as an assistant professor in 1994, becoming an associate professor in 1997 and a professor in 1999. In 1998, he returned to Israel, resigning from Princeton in 2000. His research focuses on nonlinear optics, solitons, subwavelength imaging, and quantum electronics. He is a fellow of the Optical Society of America (OSA) and the American Physical Society (APS). He won numerous awards, among them the Quantum Electronics Prize of the European Physics Society, the Max Born Award of the OSA, and the Arthur Schawlow Prize of the APS. In 2011, he was elected to the Israel Academy of Sciences and Humanities, and in 2014 he won the Israel Prize in Physics. However, beyond his personal achievements, he takes pride in the success of his students and postdocs; 17 are now professors and others are leaders in industry.

REFERENCES

[1] D. Sayre. (1952, Nov.). Some implications of a theorem due to Shannon. *Acta Crystallogr.* [Online]. 5(6), pp. 843–843. Available: http://scripts.iucr.org/cgi-bin/paper?a00763

[2] J. R. Fienup. (1978, July). Reconstruction of an object from the modulus of its Fourier transform. *Opt. Lett.* [Online]. 3(1), pp. 27–29. Available: http://ol.osa.org/abstract.cfm?URI=ol-3-1-27

[3] J. Miao, P. Charalambous, J. Kirz and D. Sayre, "Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature*, vol. 400, no. 6742, pp. 342–344, 1999.

[4] M. R. Howells, T. Beetz, H. N. Chapman, C. Cu, J. M. Holton, C. J. Jacobsen, J. Kirz, E. Lima, S. Marchesini, H. Miao, D. Sayre, D. A. Shapiro, J. C. H. Spence, and D. Starodub, (2009). An assessment of the resolution limitation due to radiation-damage in X-ray diffraction microscopy. *J. Electron Spectrosc. Relat. Phenom.* [Online]. 170(1–3), pp. 4–12. Available: http://www.sciencedirect.com/science/article/pii/S0368204808001424

[5] J. C. Solem and G. C. Baldwin, "Microholography of living organisms," *Science*, vol. 218, no. 4569, pp. 229–235, 1982.

[6] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, "Potential for biomolecular imaging with femtosecond X-ray pulses," *Nature*, vol. 406, no. 6797, pp. 752–757, 2000.

[7] J. Spence, U. Weierstall, and M. Howells, "Coherence and sampling requirements for diffractive imaging," *Ultramicroscopy*, vol. 101, no. 2, pp. 149–152, 2004.

[8] V. Elser, "Solution of the crystallographic phase problem by iterated projections," *Acta Crystallogr. Sect. A*, vol. 59, no. 3, pp. 201–209, 2003.

[9] S. Marchesini, "Invited article: A unified evaluation of iterative projection algorithms for phase retrieval," *Rev. Sci. Instrum.*, vol. 78, no. 1, pp. 011 301–011 301, 2007.

[10] H. H. Bauschke, P. L. Combettes, and D. R. Luke, "Hybrid projection–reflection method for phase retrieval," *JOSA A*, vol. 20, no. 6, pp. 1025–1034, 2003.

[11] D. R. Luke, "Relaxed averaged alternating reflections for diffraction imaging," *Inverse Probl.*, vol. 21, no. 1, p. 37, 2005.

[12] J. A. Rodriguez, R. Xu, C.-C. Chen, Y. Zou, and J. Miao, "Oversampling smoothness: An effective algorithm for phase retrieval of noisy diffraction intensities," *J. Appl. Crystallogr.*, vol. 46, no. 2, pp. 312–318, 2013.

[13] I. K. Robinson, I. A. Vartanyants, G. J. Williams, M. A. Pfeifer, and J. A. Pitney, "Reconstruction of the shapes of gold nanocrystals using coherent X-ray diffraction," *Phys. Rev. Lett.*, vol. 87, no. 19, p. 195505, 2001.

[14] B. Abbey, K. A. Nugent, G. J. Williams, J. N. Clark, A. G. Peele, M. A. Pfeifer, M. D. Jonge, and I. McNulty, "Keyhole coherent diffractive imaging," *Nat. Phys.*, vol. 4, no. 5, pp. 394–398, 2008.

[15] D. Nam, J. Park, M. Gallagher-Jones et al., "Imaging fully hydrated whole cells by coherent X-ray diffraction microscopy," *Phys. Rev. Lett.*, vol. 110, no. 9, pp. 098 103–098 103, 2013.

[16] H. N. Chapman, A. Barty, M. J. Bogan, S. Boutet, M. Frank, S. P. Hau-Riege, S. Marchesini, B. W. Woods, S. C. Bajt, W. H. Benner, R. A. London, E. Plönjes, M. Kuhlmann, R. Treusch, S. Düsterer, T. Tschentscher, J. R. Schneider, E. Spiller, T. Möller, C. Bostedt, M. Hoener, D. A. Shapiro, K. O. Hodgson, D. van der Spoel, F. Burmeister, M. Bergh, C. Caleman, G. Huldt, M. M. Seibert, F. R. N. C. Maia, R. W. Lee, A. Szöke, N. Timneanu and J. Hajdu, "Femtosecond diffractive imaging with a soft-X-ray free-electron laser," *Nat. Phys.*, vol. 2, no. 12, pp. 839–843, 2006.

[17] M. M. Seibert, T. Ekeberg, F. R. N. C. Maia, M. Svenda, J. Andreasson, O. Jönsson, D. Odić, B. Iwan, A. Rocker, D. Westphal, M. Hantke, D. P. DePonte, A. Barty, J. Schulz, L. Gumprecht, N. Coppola, A. Aquila, M. Liang, T. A. White, A. Martin, C. Caleman, S. Stern, C. Abergel, V. Seltzer, J.-M. Claverie, C. Bostedt, J. D. Bozek, S. Boutet, A. A. Miahnahri, M. Messerschmidt, J. Krzywinski, G. Williams, K. O. Hodgson, M. J. Bogan, C. Y. Hampton, R. G. Sierra, D. Starodub, I. Andersson, S. Bajt, M. Barthelmess, J. C. H. Spence, P. Fromme, U. Weierstall, R. Kirian, M. Hunter, R. B. Doak, S. Marchesini, S. P. Hau-Riege, M. Frank, R. L. Shoeman, L. Lomb, S. W. Epp, R. Hartmann, D. Rolles, A. Rudenko, C. Schmidt, L. Foucar, N. Kimmel, P. Holl, B. Rudek, B. Erk, A. Hömke, C. Reich, D. Pietschner, G. Weidenspointner, L. Strüder, G. Hauser, H. Gorke, J. Ullrich, I. Schlichting, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.-U. Kühnel, R. Andritschke, C.-D. Schröter, F. Krasniqi, M. Bott, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, H. N. Chapman and J. Hajdu, (2011, Feb.). Single mimivirus particles intercepted and imaged with an X-ray laser. *Nature*. [Online]. 470(7332), pp. 78–81. Available: http://dx.doi.org/10.1038/nature09748.

[18] R. L. Sandberg, A. Paul, D. A. Raymondson, S. Hädrich, D. M. Gaudiosi, J. Holtsnider, R. I. Tobey, O. Cohen, M. M. Murnane, H. C. Kapteyn, C. Song, J. Miao, Y. Liu, and F. Salmassi. (2007, Aug.). Lensless diffractive imaging using tabletop coherent high-harmonic soft-X-ray beams. *Phys. Rev. Lett.* [Online]. 99(9), p. 098103. Available: http://link.aps.org/doi/10.1103/PhysRevLett.99.098103

[19] B. Chen, R. A. Dilanian, S. Teichmann, B. Abbey, A. G. Peele, G. J. Williams, P. Hannaford, L. V. Dao, H. M. Quiney, and K. A. Nugent, "Multiple wavelength diffractive imaging," *Phys. Rev. A*, vol. 79, no. 2, p. 23809, 2009.

[20] M. D. Seaberg, D. E. Adams, E. L. Townsend, D. A. Raymondson, W. F. Schlotter, Y. Liu, C. S. Menoni, L. Rong, C.-C. Chen, J. Miao, H. C. Kapteyn, and M. M. Murnane. (2011, Nov.). Ultrahigh 22 nm resolution coherent diffractive imaging using a desktop 13 nm high harmonic source. *Opt. Express*. [Online]. 19(23), pp. 22 470–22 479. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-19-23-22470

[21] R. L. Sandberg, C. Song, P. W. Wachulak, D. A. Raymondson, A. Paul, B. Amirbekian, E. Lee, A. E. Sakdinawat, C. La-O-Vorakiat, M. C. Marconi, C. S. Menoni, M. M. Murnane, J. J. Rocca, H. C. Kapteyn, and J. Miao, "High numerical aperture tabletop soft X-ray diffraction microscopy with 70-nm resolution," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 105, no. 1, pp. 24–27, 2008.

[22] J. Bertolotti, E. G. van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, "Non-invasive imaging through opaque scattering layers," *Nature*, vol. 491, no. 7423, pp. 232–234, 2012.

[23] J. M. Zuo, I. Vartanyants, M. Gao, R. Zhang, and L. A. Nagahara, "Atomic resolution imaging of a carbon nanotube from diffraction intensities," *Science*, vol. 300, no. 5624, pp. 1419–1421, 2003.

[24] C. T. Putkunz, A. J. D'Alfonso, A. J. Morgan, M. Weyland, C. Dwyer, L. Bourgeois, J. Etheridge, A. Roberts, R. E. Scholten, K. A. Nugent, and L. J. Allen, "Atom-scale ptychographic electron diffractive imaging of boron nitride cones," *Phys. Rev. Lett.*, vol. 108, no. 7, p. 073901, 2012.

[25] P. Thibault and V. Elser, "X-ray diffraction microscopy," *Annu. Rev. Condens. Matter Phys.*, vol. 1, no. 1, pp. 237–255, 2010.

[26] K. A. Nugent, "Coherent methods in the X-ray sciences," *Adv. Phys.*, vol. 59, no. 1, pp. 1–99, 2010.

[27] C. Song, H. Jiang, A. Mancuso, B. Amirbekian, L. Peng, R. Sun, S. S. Shah, Z. H. Zhou, T. Ishikawa, and J. Miao, "Quantitative imaging of single, unstained viruses with coherent x rays," *Phys. Rev. Lett.*, vol. 101, no. 15, p. 158101, 2008.

[28] N. D. Loh, C. Y. Hampton, A. V. Martin, D. Starodub, R. G. Sierra, A. Barty, A. Aquila, J. Schulz, L. Lomb, J. Steinbrener, R. L. Shoeman, S. Kassemeyer, C. Bostedt, J. Bozek, S. W. Epp, B. Erk, R. Hartmann, D. Rolles, A. Rudenko, B. Rudek, L. Foucar, N. Kimmel, G. Weidenspointner, G. Hauser, P. Holl, E. Pedersoli, M. Liang, M. S. Hunter, L. Gumprecht, N. Coppola, C. Wunderer, H. Graafsma, F. R. N. C. Maia, T. Ekeberg, M. Hantke, H. Fleckenstein, H. Hirsemann, K. Nass, T. A. White, H. J. Tobias, G. R. Farquar, W. H. Benner, S. P. Hau-Riege, C. Reich, A. Hartmann, H. Soltau, S. Marchesini, S. Bajt, M. Barthelmess, P. Bucksbaum, K. O. Hodgson, L. Strüder, J. Ullrich, M. Frank, I. Schlichting, H. N. Chapman and M. J. Bogan. (2012, June). Fractal morphology, imaging and mass spectrometry of single aerosol particles in flight. *Nature*. [Online]. 486(7404), pp. 513–517. Available: http://www.nature.com/nature/journal/v486/n7404/abs/nature11222.html

[29] R. A. Gonsalves. (1982). Phase retrieval and diversity in adaptive optics. *Opt. Eng.* [Online]. 21(5), pp. 215–829. Available: http://dx.doi.org/10.1117/12.7972989

[30] J. E. Krist and C. J. Burrows, "Phase-retrieval analysis of pre-and post-repair hubble space telescope images," *Appl. Opt.*, vol. 34, no. 22, pp. 4951–4964, 1995.

[31] D. R. Luke, J. V. Burke, and R. G. Lyon, "Optical wavefront reconstruction: Theory and numerical methods," *SIAM Rev.*, vol. 44, no. 2, pp. 169–224, 2002.

[32] A. Labeyrie, "Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images," *Astron. Astrophys.*, vol. 6, no. 1, pp. 85–87, 1970.

[33] J. C. Dainty and J. R. Fienup, "Phase retrieval and image reconstruction for astronomy," in *Image Recovery: Theory and Application*, H. Stark, Ed. San Diego, CA: Academic Press, 1987, pp. 231–275.

[34] C. Lynds, S. Worden, and J. W. Harvey, "Digital image reconstruction applied to alpha orionis," *Astrophys. J.*, vol. 207, no. 1, pp. 174–180, July 1976.

[35] A. W. Lohmann, G. Weigelt, and B. Wirnitzer, "Speckle masking in astronomy: Triple correlation theory and applications," *Appl. Opt.*, vol. 22, no. 24, pp. 4028–4037, 1983.

[36] F. Roddier, *Adaptive Optics in Astronomy*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[37] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*. Hoboken, NJ: Wiley, Mar. 2007.

[38] E. J. Candes, T. Strohmer, and V. Voroninski, "PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming," *Commun. Pure Appl. Math.*, vol. 66, no. 8, pp. 1241–1274, 2013.

[39] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM J. Imag. Sci.*, vol. 6, no. 1, pp. 199–225, 2013.

[40] R. Balan, P. Casazza, and D. Edidin, "On signal reconstruction without phase," *Appl. Comput. Harm. Anal.*, vol. 20, no. 3, pp. 345–356, 2006.

[41] M. L. Moravec, J. K. Romberg, and R. G. Baraniuk, "Compressive phase retrieval," *Proc. SPIE*, vol. 6701, Wavelets XII, p. 670120, Sept. 2007.

[42] A. Fannjiang, "Absolute uniqueness of phase retrieval with random illumination," *Inverse Probl.*, vol. 28, no. 7, p. 075008, 2012.

[43] A. S. Bandeira, J. Cahill, D. G. Mixon, and A. A. Nelson, "Saving phase: Injectivity and stability for phase retrieval," *Appl. Computat. Harm. Anal.,* vol. 37, no. 1, pp. 106–125, 2014.

[44] A. Beck and Y. C. Eldar. (2013). Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. *SIAM J. Optim.* [Online]. 23(3), pp. 1480–1509. Available: http://epubs.siam.org/doi/abs/10.1137/120869778

[45] A. Szameit, Y. Shechtman, E. Osherovich, E. Bullkich, P. Sidorenko, H. Dana, S. Steiner, E. B. Kley, S. Gazit, T. Cohen-Hyams, S. Shoham, M. Zibulevsky, I. Yavneh, Y. C. Eldar, O. Cohen, and M. Segev. (2012). Sparsity-based single-shot subwavelength coherent diffractive imaging. *Nat. Mater.* [Online]. 11(5), pp. 455–459. Available: http://www.nature.com/nmat/journal/v11/n5/full/nmat3289.html

[46] Y. Shechtman, Y. C. Eldar, A. Szameit, and M. Segev. (2011, Aug.). Sparsity based sub-wavelength imaging with partially incoherent light via quadratic compressed sensing," *Opt. Express.* [Online]. 19(16), pp. 14 807–14 822. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-19-16-14807

[47] H. Ohlsson, A. Y. Yang, R. Dong, and S. S. Sastry, "Compressive phase retrieval from squared output measurements via semidefinite programming," presented at the 16th IFAC Symposium on System Identification, SYSID 2012.

[48] S. Bahmani, P. Boufounos, and B. Raj, "Greedy sparsity-constrained optimization," in *2011 Conf. Rec. Forty Fifth Asilomar Conf. Signals, Systems and Computers (ASILOMAR)*, 2011, pp. 1148–1152.

[49] K. Jaganathan, S. Oymak, and B. Hassibi, "Recovery of sparse 1-d signals from the magnitudes of their fourier transform," in *Proc. 2012 IEEE Int. Symp. Information Theory (ISIT)*, 2012, pp. 1473–1477.

[50] Y. C. Eldar and S. Mendelson. (2013). Phase retrieval: Stability and recovery guarantees. *Appl. Comput. Harm. Anal.* [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1063520313000717

[51] Y. Shechtman, A. Beck, and Y. C. Eldar, "GESPAR: Efficient phase retrieval of sparse signals," *IEEE Trans. Signal Processing*, vol. 62, no. 1–4, pp. 928–938, 2014.

[52] Y. Shechtman, Y. C. Eldar, O. Cohen, and M. Segev. (2013, Mar.). Efficient coherent diffractive imaging for sparsely varying objects. *Opt. Express.* [Online]. 21(5), pp. 6327–6338. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-21-5-6327.

[53] Y. Shechtman, E. Small, Y. Lahini, M. Verbin, Y. C. Eldar, Y. Silberberg, and M. Segev. (2013, Oct.). Sparsity-based superresolution and phase-retrieval in waveguide arrays. *Opt. Express.* [Online]. 21(20), pp. 24 015–24 024. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-21-20-24015

[54] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications.* Cambridge, U.K.: Cambridge Univ. Press, 2012.

[55] P. Emma, R. Akre, J. Arthur, R. Bionta, C. Bostedt, J. Bozek, A. Brachmann, P. Bucksbaum, R. Coffee, F.-J. Decker, Y. Ding, D. Dowell, S. Edstrom, A. Fisher, J. Frisch, S. Gilevich, J. Hastings, G. Hays, Ph. Hering, Z. Huang, R. Iverson, H. Loos, M. Messerschmidt, A. Miahnahri, S. Moeller, H.-D. Nuhn, G. Pile, D. Ratner, J. Rzepiela, D. Schultz, T. Smith, P. Stefan, H. Tompkins, J. Turner, J. Welch, W. White, J. Wu, G. Yocky and J. Galayda, "First lasing and operation of an Ångstrom-wavelength free-electron laser," *Nat. Photon.*, vol. 4, no. 9, pp. 641–647, 2010.

[56] T. Ishikawa, H. Aoyagi, T. Asaka, Y. Asano, N. Azumi, T. Bizen, H. Ego, K. Fukami, T. Fukui, Y. Furukawa, S. Goto, H. Hanaki, T. Hara, T. Hasegawa, T. Hatsui, A. Higashiya, T. Hirono, N. Hosoda, M. Ishii, T. Inagaki, Y. Inubushi, T. Itoga, Y. Joti, M. Kago, T. Kameshima, H. Kimura, Y. Kirihara, A. Kiyomichi, T. Kobayashi, C. Kondo, T. Kudo, H. Maesaka, X. M. Maréchal, T. Masuda, S. Matsubara, T. Matsumoto, T. Matsushita, S. Matsui, M. Nagasono, N. Nariyama, H. Ohashi, T. Ohata, T. Ohshima, S. Ono, Y. Otake, C. Saji, T. Sakurai, T. Sato, K. Sawada, T. Seike, K. Shirasawa, T. Sugimoto, S. Suzuki, S. Takahashi, H. Takebe, K. Takeshita, K. Tamasaku, H. Tanaka, R. Tanaka, T. Tanaka, T. Togashi, K. Togawa, A. Tokuhisa, H. Tomizawa, K. Tono, S. Wu, M. Yabashi, M. Yamaga, A. Yamashita, K. Yanagida, C. Zhang, T. Shintake, H. Kitamura, and N. Kumagai, "A compact X-ray free-electron laser emitting in the sub-angstrom region," *Nat. Photon.*, vol. 6, no. 8, pp. 540–544, 2012.

[57] I. Waldspurger, A. d'Aspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Math. Program.*, Ser. A, pp. 1–35, Dec. 2013.

[58] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," *Adv. Neural Inform. Process. Syst.*, pp. 2796–2804, 2013.

[59] X. Li and V. Voroninski, "Sparse signal recovery from quadratic measurements via convex programming," *SIAM J. Math. Anal.*, vol. 45, no. 5, pp. 3019–3033, 2013.

[60] G., David, F. Krahmer, and R. Kueng. "A partial derandomization of Phase Lift using spherical designs." arXiv preprint arXiv:1310.2267 (2013).

[61] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proc. IEEE*, vol. 69, no. 5, pp. 529–541, 1981.

[62] E. Hofstetter, "Construction of time-limited functions with specified autocorrelation functions," *IEEE Trans. Inform. Theory*, vol. 10, no. 2, pp. 119–126, 1964.

[63] Y. M. Bruck and L. Sodin, "On the ambiguity of the image reconstruction problem," *Opt. Commun.*, vol. 30, no. 3, pp. 304–308, 1979.

[64] M. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform," *IEEE Trans. Acoust., Speech Signal Processing*, vol. 30, no. 2, pp. 140–154, 1982.

[65] R. H. T. Bates, "Fourier phase problems are uniquely solvable in more than one dimension. I: Underlying theory," *Optik*, vol. 61, no. 3, pp. 247–262, 1982.

[66] J. Miao, D. Sayre, and H. Chapman, "Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects," *JOSA A*, vol. 15, no. 6, pp. 1662–1669, 1998.

[67] P. L. Van Hove, M. H. Hayes, J. S. Lim, and A. V. Oppenheim, "Signal reconstruction from signed fourier transform magnitude," *IEEE Trans. Acoust., Speech Signal Processing*, vol. 31, no. 5, pp. 1286–1293, 1983.

[68] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing.* New York: Springer-Verlag, 2010.

[69] J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, "Phase retrieval for sparse signals: Uniqueness conditions," Preprints, arXiv:1308.3058 (2013).

[70] H. Ohlsson and Y. C. Eldar, "On conditions for uniqueness in sparse phase retrieval," Preprints, arXiv:1308.5447 (2013).

[71] K. Jaganathan, S. Oymak, and B. Hassibi. (2013, Nov.). Sparse phase retrieval: Uniqueness guarantees and recovery algorithms. arXiv:1311.2745 [cs, math]. [Online]. Available: http://arxiv.org/abs/1311.2745

[72] E. Osherovich, M. Zibulevsky, and I. Yavneh, "Approximate Fourier phase information in the phase retrieval problem: What it gives and how to use it," *JOSA A*, vol. 28, no. 10, pp. 2124–2131, 2011.

[73] E. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval from masked Fourier transforms," Preprints, arXiv:1310.3240 (2013).

[74] R. Gerchberg and W. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, p. 237, 1972.

[75] J. Fienup, "Phase retrieval algorithms: A comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, 1982.

[76] D. Misell, "An examination of an iterative method for the solution of the phase problem in optics and electron optics: I. test calculations," *J. Phys. D: Appl. Phys.*, vol. 6, no. 18, p. 2200, 1973.

[77] A. Huiser and H. Ferwerda, "The problem of phase retrieval in light and electron microscopy of strong objects: Ii. On the uniqueness and stability of object reconstruction procedures using two defocused images," *J. Mod. Opt.*, vol. 23, no. 6, pp. 445–456, 1976.

[78] W. Hoppe and G. Strube, "Beugung in inhomogenen primarstrahlenwellenfeld. II. lichtoptische analogieversuche zur phasenmessung von gitterinterferenzen," *Acta Crystallogr. Sect. A*, vol. 25, no. 4, pp. 502–507, 1969.

[79] J. Rodenburg, "Ptychography and related diffractive imaging methods," *Adv. Imag. Electron Phys.*, vol. 150, pp. 87–184, 2008.

[80] M. Dierolf, A. Menzel, P. Thibault, P. Schneider, C. M. Kewish, R. Wepf, O. Bunk and F. Pfeiffer, "Ptychographic X-ray computed tomography at the nanoscale," *Nature*, vol. 467, no. 7314, pp. 436–439, 2010.

[81] M. Reed Teague, "Deterministic phase retrieval: A green's function solution," *JOSA*, vol. 73, no. 11, pp. 1434–1441, 1983.

[82] N. Streibl. (1984). Phase imaging by the transport equation of intensity. *Opt. Commun.* [Online]. 49(1), pp. 6–10. Available: http://www.sciencedirect.com/science/article/pii/0030401884900798

[83] T. E. Gureyev and K. A. Nugent, "Phase retrieval with the transport-of-intensity equation. ii. orthogonal series solution for nonuniform illumination," *JOSA A*, vol. 13, no. 8, pp. 1670–1682, 1996.

[84] J. R. Fienup. (1987, Jan.). Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint. *J. Opt. Soc. Am. A.* [Online]. 4(1), pp. 118–123. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-4-1-118

[85] J. Miao, Y. Nishino, Y. Kohmura, B. Johnson, C. Song, S. H. Risbud, and T. Ishikawa, "Quantitative image reconstruction of GaN quantum dots from oversampled diffraction intensities alone," *Phys. Rev. Lett.*, vol. 95, no. 8, p. 085503, 2005.

[86] H. Jiang, C. Song, C.-C. Chen, R. Xu, K. S. Raines, B. P. Fahimian, C.-H. Lu, T.-K. Lee, A. Nakashima, J. Urano, T. Ishikawa, F. Tamanoi, and J. Miao, "Quantitative 3D imaging of whole, unstained cells by using X-ray diffraction microscopy," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 107, no. 25, pp. 11 234–11 239, 2010.

[87] C.-C. Chen, J. Miao, C. W. Wang, and T. K. Lee, "Application of optimization technique to noncrystalline X-ray diffraction microscopy: Guided hybrid input-output method," *Phys. Rev., Ser. B*, vol. 76, no. 6, p. 064113, 2007.

[88] A. V. Martin, F. Wang, N. D. Loh, T. Ekeberg, F. R. N. C. Maia, M. Hantke, G. van der Schot, C. Y. Hampton, R. G. Sierra, A. Aquila, S. Bajt, M. Barthelmess, C. Bostedt, J. D. Bozek, N. Coppola, S. W. Epp, B. Erk, H. Fleckenstein, L. Foucar, M. Frank, H. Graafsma, L. Gumprecht, A. Hartmann, R. Hartmann, G. Hauser, H. Hirsemann, P. Holl, S. Kassemeyer, N. Kimmel, M. Liang, L. Lomb, S. Marchesini, K. Nass, E. Pedersoli, C. Reich, D. Rolles, B. Rudek, A. Rudenko, J. Schulz, R. L. Shoeman, H. Soltau, D. Starodub, J. Steinbrener, F. Stellato, L. Strüder, J. Ullrich, G. Weidenspointner, T. A. White, C. B. Wunderer, A. Barty, I. Schlichting, M. J. Bogan, and H. N. Chapman. (2012, July). "Noise-robust coherent diffractive imaging with a single diffraction pattern," *Opt. Express.* [Online]. 2015), pp. 16,650–16,661. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-15-16650.

[89] D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1 - theory," *IEEE Trans. Medical Imaging*, vol. 1, no. 2, pp. 81–94, 1982.

[90] A. Levi and H. Stark, "Signal restoration from phase by projections onto convex sets," *JOSA*, vol. 73, no. 6, pp. 810–822, 1983.

[91] A. Levi and H. Stark, "Image restoration by the method of generalized projections with application to restoration from magnitude," *JOSA A*, vol. 1, no. 9, pp. 932–943, 1984.

[92] H. Bauschke, P. Combettes, and D. Luke, "Phase retrieval, error reduction algorithm, and Fienup variants: A view from convex optimization," *JOSA A*, vol. 19, no. 7, pp. 1334–1345, 2002.

[93] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "A block coordinate variable metric forward-backward algorithm," *Optim. Online*, 2013. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00945918/

[94] J. Douglas and H. Rachford, "On the numerical solution of heat conduction problems in two and three space variables," *Trans. Amer. Math. Soc.*, vol. 82, no. 2, pp. 421–439, 1956.

[95] P.-L. Lions and B. Mercier, "Splitting algorithms for the sum of two nonlinear operators," *SIAM J. Numer. Anal.*, vol. 16, no. 6, pp. 964–979, 1979.

[96] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, 1996.

[97] M. Fazel, H. Hindi, and S. P. Boyd, "Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices," in *Proc. 2003 IEEE American Control Conf.*, vol. 3, pp. 2156–2162.

[98] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *J. ACM*, vol. 42, no. 6, pp. 1115–1145, 1995.

[99] F. Roddier, "Wavefront sensing and the irradiance transport equation," *Appl. Opt.*, vol. 29, no. 10, pp. 1402–1403, 1990.

[100] L. Waller, L. Tian, and G. Barbastathis, "Transport of intensity phase-amplitude imaging with higher order intensity derivatives," *Opt. Express*, vol. 18, no. 12, pp. 12 552–12 561, 2010.

[101] D. Paganin and K. A. Nugent, "Noninterferometric phase imaging with partially coherent light," *Phys. Rev. Lett.*, vol. 80, no. 12, p. 2586, 1998.

[102] D. Paganin, A. Barty, P. McMahon et al., "Quantitative phase-amplitude microscopy. III. The effects of noise," *J. Microsc.*, vol. 214, no. 1, pp. 51–61, 2004.

[103] S. Mukherjee and C. Seelamantula, "An iterative algorithm for phase retrieval with sparsity constraints: Application to frequency domain optical coherence tomography," in *Proc. 2012 IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 553– 556.

[104] G. Oszlányi and A. Suto, "Ab initio structure solution by charge flipping," *Acta Crystallogr., Sect. A*, vol. 60, no. 2, pp. 134–141, 2004.

[105] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[106] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[107] D. L. Donoho and M. Elad. (2003, Mar.). Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization. *Proc. Natl. Acad. Sci. U.S.A.* [Online]. 100(5), pp. 2197–2202. Available: http://www.pnas.org/content/100/5/2197

[108] Y. C. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *1993 Conf. Rec. 27th Asilomar Conf. Signals, Systems and Computers,* IEEE, 1993, pp. 40–44.

[109] S. S. Chen, D. L. Donoho, and M. A. Saunders. (1998, Jan.). Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.* [Online]. 20(1), pp. 33–61. Available: http://epubs.siam.org/doi/abs/10.1137/S1064827596304010

[110] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[111] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4053–4085, 2011.

[112] K. Jaganathan, S. Oymak, and B. Hassibi, "Recovery of sparse 1-D signals from the magnitudes of their Fourier transform," in *2012 IEEE Int. Symp. Information Theory Proc. (ISIT)*, 2012, pp. 1473–1477.

[113] A. Beck and Y. C. Eldar. (2012, Mar.). Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. *arXiv*:1203.4580. [Online]. Available: http://arxiv.org/abs/1203.4580

[114] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. New York: Dover, 1998.

[115] D. P. Bertsekas, "Nonlinear programming," 1999.

[116] P. Sidorenko, A. Fleischer, Y. Shechtman, Y. C. Eldar, M. Segev, and O. Cohen,, "Sparsity-based superresolution coherent diffractive imaging of (practically) 1d images using extreme UV radiation," in *Proc. CLEO: 2013*. Optical Society of America, 2013, p. QF1C.7

[117] C. Song, D. Ramunno-Johnson, Y. Nishino, Y. Kohmura, T. Ishikawa, C.-C. Chen, T.-K. Lee, and J. Miao, "Phase retrieval from exactly oversampled diffraction intensity through deconvolution," *Phys. Rev. B*, vol. 75, no. 1, p. 012102, 2007.

[118] L. Whitehead, G. J. Williams, H. M. Quiney, D. J. Vine, R. A. Dilanian, S. Flewett, K. A. Nugent, A. G. Peele, E. Balaur, and I. McNulty,, "Diffractive imaging using partially coherent x rays," *Phys. Rev. Lett.*, vol. 103, no. 24, p. 243902, 2009.

[119] G. J. Williams, M. A. Pfeifer, I. A. Vartanyants, and I. K. Robinson, "Three-dimensional imaging of microstructure in au nanocrystals," *Phys. Rev. Lett.*, vol. 90, no. 17, p. 175501, 2003.

[120] C. Song, H. Jiang, A. Mancuso, B. Amirbekian, Li Peng, R. Sun, S. S. Shah, Z. H. Zhou, T. Ishikawa, and J. Miao, "Quantitative imaging of single, unstained viruses with coherent x rays," *Phys. Rev. Lett.*, vol. 101, no. 15, p. 158101, 2008.

[121] N. D. Loh, M. J. Bogan, V. Elser, A. Barty, S. Boutet, S. Bajt, J. Hajdu, T. Ekeberg, F. R. N. C. Maia, J. Schulz, M. M. Seibert, B. Iwan, N. Timneanu, S. Marchesini, I. Schlichting, R. L. Shoeman, L. Lomb, M. Frank, M. Liang, and H. N. Chapman, "Cryptotomography: Reconstructing 3D Fourier intensities from randomly oriented single-shot diffraction patterns," *Phys. Rev. Lett.*, vol. 104, no. 22, p. 225501, 2010.

[122] L. Strüder, S. Epp, D. Rolles, R. Hartmann, P. Holl, G. Lutz, H. Soltau, R. Eckart, C. Reich, K. Heinzinger, C. Thamm, A. Rudenko, F. Krasniqi, K.-U. Kühnel, C. Bauer, C.-D. Schröter, R. Moshammer, S. Techert, D. Miessner, M. Porro, O. Hälker, N. Meidinger, N. Kimmel, R. Andritschke, F. Schopper, G. Weidenspointner, A. Ziegler, D. Pietschner, S. Herrmann, U. Pietsch, A. Walenta, W. Leitenberger, C. Bostedt, T. Möller, D. Rupp, M. Adolph H Graafsma, H. Hirsemann, K. Gärtner, R. Richter, L. Foucar, R. L. Shoeman, I. Schlichting, and J. Ullrich, "Large-format, high-speed, X-ray PNCCDs combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources," *Nucl. Instrum. Meth. Phys. Res. Sect. A*, vol. 614, no. 3, pp. 483–496, 2010.

[123] J. Bozek, "Amo instrumentation for the lcls X-ray fel," *Eur. Phys. J. Special Top.*, vol. 169, no. 1, pp. 129–132, 2009.

[124] S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. H. Spence. (2003, Oct.). X-ray image reconstruction from a diffraction pattern alone. *Phys. Rev. B*. [Online]. vol. 68, p. 140101. Available: http://link.aps.org/doi/10.1103/PhysRevB.68.140101

[125] H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C.H. Spence, U. Weierstall, T. Beetz, C. Jacobsen, and D. Shapiro, "High-resolution ab initio three-dimensional X-ray diffraction microscopy," *JOSA A*, vol. 23, no. 5, pp. 1179–1200, 2006.

[126] R. L. Sandberg, Z. Huang, R. Xu, J. A. Rodriguez, and J. Miao, "Studies of materials at the nanometer scale using coherent X-ray diffraction imaging," *JOM*, vol. 65, no. 9, pp. 1208–1220, 2013.

[127] L.-M. Stadler, C. Gutt, T. Autenrieth, O. Leupold, S. Rehbein, Y. Chushkin, and G. Grübel. (2008, June). Hard x ray holographic diffraction imaging. *Phys. Rev. Lett.* [Online]. 100(24), p. 245503. Available: http://link.aps.org/doi/10.1103/PhysRevLett.100.245503

[128] G. J. Williams, H. M. Quiney, B. B. Dhal, C. Q. Tran, K. A. Nugent, A. G. Peele, D. Paterson, and M. D. de Jonge, "Fresnel coherent diffractive imaging," *Phys. Rev. Lett.*, vol. 97, no. 2, p. 025506, 2006.

[129] B. Abbey, L. W. Whitehead, H. M. Quiney, D. J. Vine, G. A. Cadenazzi, C. A. Henderson, K. A. Nugent, E. Balaur, C. T. Putkunz, A. G. Peele, G. J. Williams, and I. McNulty, "Lensless imaging using broadband X-ray sources," *Nat. Photon.*, vol. 5, no. 7, pp. 420–424, 2011.

[130] S. Witte, V. T. Tenner, D. W. Noom, and K. SE Eikema, "Ultra-broadband extreme-ultraviolet lensless imaging of extended complex structures," Preprints, arXiv:1302.6064 (2013).

[131] H. N. Chapman, "Phase-retrieval X-ray microscopy by Wigner-distribution deconvolution," *Ultramicroscopy*, vol. 66, no. 3, pp. 153–172, 1996.

[132] P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer, "High-resolution scanning X-ray diffraction microscopy," *Science*, vol. 321, no. 5887, pp. 379–382, 2008.

[133] I. Peterson, B. Abbey, C.T. Putkunz, D.J. Vine, G.A. van Riessen, G.A. Cadenazzi, E. Balaur, R. Ryan, H.M. Quiney, I. McNulty, A.G. Peele, and K.A. Nugent, "Nanoscale fresnel coherent diffraction imaging tomography using ptychography," *Opt. Express*, vol. 20, no. 22, pp. 24 678–24 685, 2012.

[134] O. Shapira, A. F. Abouraddy, J. D. Joannopoulos, and Y. Fink, "Complete modal decomposition for optical waveguides," *Phys. Rev. Lett.*, vol. 94, no. 14, p. 143902, 2005.

[135] I. Robinson and R. Harder, "Coherent X-ray diffraction imaging of strain at the nanoscale," *Nat. Mater.*, vol. 8, no. 4, pp. 291–298, 2009.

[136] M. C. Newton, S. J. Leake, R. Harder, and I. K. Robinson, "Three-dimensional imaging of strain in a single zno nanorod," *Nat. Mater.*, vol. 9, no. 2, pp. 120–124, 2009.

[137] W. Yang, X. Huang, R. Harder et al. (2013, Apr.). Coherent diffraction imaging of nanoscale strain evolution in a single crystal under high pressure. *Nat Commun.* [Online]. vol. 4, p. 1680. Available: http://dx.doi.org/10.1038/ncomms2661

[138] M. Holt, R. Harder, R. Winarski, and V. Rose, "Nanoscale hard X-ray microscopy methods for materials studies," *Annu. Rev. Mater. Res.*, vol. 43, no. 1, pp. 183–211, 2013.

[139] D. Gabor, "A new microscopic principle," *Nature*, vol. 161, no. 4098, pp. 777–778, 1948.

[140] I. McNulty, J. Kirz, C. Jacobsen, E. H. Anderson, M. R. Howells, and D. P. Kern, "High-resolution imaging by Fourier transform X-ray holography," *Science*, vol. 256, no. 5059, pp. 1009–1012, 1992.

[141] S. Kikuta, S. Aoki, S. Kosaki, and K. Kohra. (1972). X-ray holography of lensless Fourier-transform type. *Opt. Commun.* [Online]. 5(2), pp. 86–89. Available: http://www.sciencedirect.com/science/article/pii/0030401872900053

[142] S. Eisebitt, J. Lüning, W. F. Schlotter, M. Lörgen, O. Hellwig, W. Eberhardt, and J. Stöhr. (2004, Dec.). Lensless imaging of magnetic nanostructures by X-ray spectro-holography. *Nature*. [Online]. 432(7019), pp. 885–888. Available: http://www.nature.com/nature/journal/v432/n7019/abs/nature03139.html

[143] T. Latychevskaia, J.-N. Longchamp, and H.-W. Fink, "When holography meets coherent diffraction imaging," *Opt. Express*, vol. 20, no. 27, pp. 28 871–28 892, 2012.

[144] T. Popmintchev, M.-C. Chen, D. Popmintchev, P. Arpin, S. Brown, S. Ališauskas, G. Andriukaitis, T. Balčiunas, O. D. Mücke, A. Pugzlys, A. Baltuška, B. Shim, S. E. Schrauth, A. Gaeta, C. Hernández-García, L. Plaja, A. Becker, A. Jaron-Becker, M. M. Murnane, and H. C. Kapteyn, "Bright coherent ultrahigh harmonics in the keV X-ray regime from mid-infrared femtosecond lasers," *Science*, vol. 336, no. 6086, pp. 1287–1291, 2012.

[145] J. Miao, T. Ishikawa, B. Johnson, E. H. Anderson, B. Lai, and K. O. Hodgson, "High resolution 3D X-ray diffraction microscopy," *Phys. Rev. Lett.*, vol. 89, p. 088303, 2002.

[SP]

Hagit Messer and Omry Sendik

A New Approach to Precipitation Monitoring

A critical survey of existing technologies and challenges

Accurate measurements of precipitation are essential for many applications, ranging from flash-flood warnings to water resource management. However, the accuracy of the existing tools is limited by various technical and practical reasons. Percipitation monitoring has traditionally been known to rely on gauges, weather radars, and satellites. Recently, a new approach has begun to be examined, the usage of commercial wireless communication networks (CWCNs), which enjoys the lack of any need for deployment procedures or costs, and which is already widely spread across countries.

The goal of this article is to present a critical survey of the existing papers and works on this topic. We emphasize the works relating this topic to multidimensional signal processing. The importance of precipitation (rain, sleet, hail, snow, and any other outcomes of the condensation of water vapor that falls by virtue of gravity) is clear to any layman. Whether it is required for the purpose of precisely measuring past precipitation quantities or for generating future predictions, monitoring such phenomena has been of interest to humankind since early biblical days.

The differences between the various types of monitoring methods are vast and become crucial when deciding which method to use, when to use it, and where to use it. The difference between the tools ranges from their measurements' precision to their spatial and temporal nature (local-/short- versus global-/long-range) and even deployment prices [1]. Such differences are the provenances of a multitude of active research fields. These include numerical techniques for reconstructing rain maps, methods for assimilating the various monitoring methods, and, naturally, the development of new and more advanced measurement systems.

The first evidence of intentional rain gauge usage dates back to the fourth century B.C. in India [2]. Yet, contemporary rain gauges (tipping bucket and electronic gauges) are still being improved. The development of designated microwave (MW) radar dates back to the late 1940s [2], [3], and the development of cheaper, more precise radar has been a work in progress ever since.

Satellite-based measurements entered the environmental monitoring turf in the 1960s. Since then, the challenge of gaining precise measurements from these satellites has been a great effort. However, lowering the prices of satellite-based monitoring systems still seems to be a distant goal, and precipitation measurements from satellites are still not frequently updated.

Recently, much interest has grown around the subject of using existing CWCNs for rain monitoring [4]–[6]. Such rain-monitoring systems benefit mainly but not solely from not needing to deploy any sensors. Making use of the existing commercial (e.g., cellular) wireless networks is the equivalent of deploying a very high density of dedicated sensors but without any extra cost. Such an amount of sensors, used for precipitation monitoring, is unprecedented and can provide high temporal and spatial resolution sensing and better area coverage as well as a diversity of measurements in given points. Moreover, by applying advanced signal processing algorithms, which exploit the diversity in the data, overcoming many of the disadvantages of the previous monitoring methods now seems realizable. Such algorithms also benefit from the recent rising interest in wireless sensor networks.

NEW MONITORING APPROACH

MICROWAVE LINKS MEASUREMENT SYSTEMS
Recently, a new approach has entered the discipline of precipitation monitoring: using MW attenuation measurements for reconstruction of rainfall fields, which was initially suggested by Giuli et al. [7], [8]. Reference [7] suggested a custom design of MW links, which was set to ensure a proper reconstruction of rain fields.

A project named Microwave Attenuation as a New Tool for Improving Stormwater Supervision Administration (MANTISSA) [9] set out to test the feasibility of using MW signals to estimate rainfall. These signals are inherently path averages since they are the result of an integrated sample of the signal along the MW's path. MANTISSA aspired to use these averaged rainfall estimates as a complement to radar data and to improve the available input data to hydrological models for forecasting the response of urban and rural drainage systems.

A novel method, suggested by Messer et al. [4] in 2006, followed by Leijnse et al. [6] in 2007, involving existing CWCNs

suggested the usage of the backhaul communication links for the sake of precipitation monitoring. In other words, Messer et al. suggested using existing cellular networks' equipment for the sake of meteorological monitoring of rainfall. This suggestion alleviated the problem of the costs of the MW-based systems by using the existing links, which changed their high deployment price to zero.

Evidently, the received signal level (RSL) strength at which each antenna receives its pair's transmitted signal may be stored. Moreover, it is indeed often stored and kept for offline inspection. Messer et al. [4] proposed the usage of these cellular networks' built-in monitoring facilities. Being a widely distributed observation network, operating in real time with minimum supervision and without additional cost [4], motivated the attempt to use these data from the CWCNs. The theoretical justification for such attempts is a power law that relates the signal attenuation to the rain rate [10]. The power law relating the attenuation to the rain rate was shown to be an approximation, which holds in convective rains and in communication systems operating in midrange frequencies (above 1 GHz and below the optical range). The exact relation between the attenuation and rain rate is given by a series relation dependent on the frequency, the temperature, and the drop size distribution (DSD). Later, Olsen et al. [10] also showed that using the approximation of

$$A = aR^b \qquad (1)$$

is good, where A is the logarithmic attenuation per kilometer $A[\text{dB/km}]$ and $R[\text{mm/h}]$ is the rain rate, and they evaluated its usage with experimental results. The $A-R$ relation is often considered completely linear, approximating the power coefficient b to 1, when operating at around 1-cm wavelengths. In the dedicated MW links, which were suggested by Giuli et al. [7], [8], the frequencies were chosen to ensure a linear $A-R$ relation.

In the system devised by Giuli et al., the geometry of the links was designed to ensure a proper reconstruction of rain maps inside an area of 400 km^2. In the CWCN system suggested by Messer et al., the links geometry was designed for any arbitrary means. Placement of communication links, as performed by network technicians, is an intricate task. The execution of this task usually balances between attempts to minimize the number of calls that will be lost due to a lack of reception and attempts to minimize the number of links to reduce network establishment costs. Such an optimization target unsurprisingly generates a completely undefined geometry of a spatial distribution of links. Figure 1 depicts the Giuli link system geometry compared with the link distribution in Israel. Such arbitrary distributions hint at the challenges that CWCN-based reconstruction induces.

In a CWCN, dedicated pairs of antennas communicate with each other to transfer various types of data (audio data, billing data, etc.). The RSL strengths at which each antenna receives its dedicated pair's transmitted signal are sampled and logged. Assuming a sufficient quantity of antennas are contained in an area of interest, and a satisfactory amount of samples are in hand, a reconstruction of the rain in this area of interest may be achieved using the RSL data. An attempt to reconstruct rainfall maps by processing the recordings of the RSLs of the CWCN
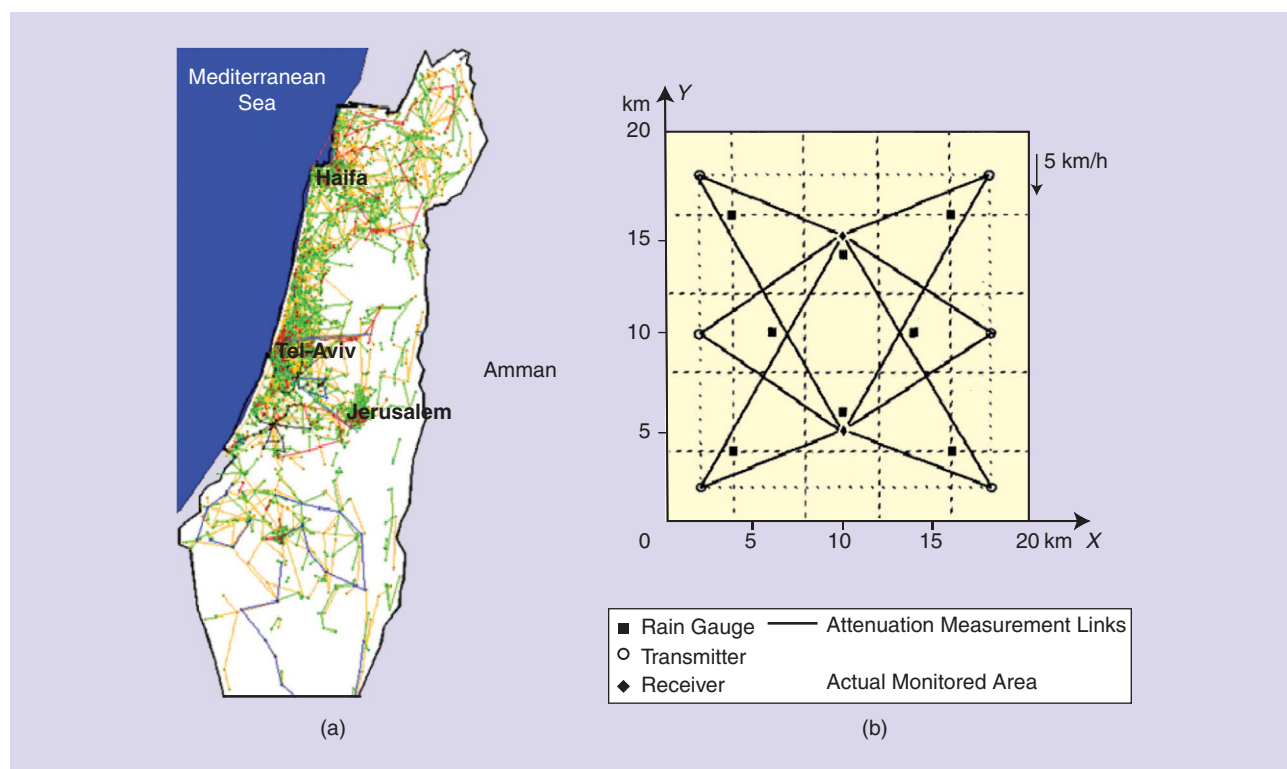
[FIG1] (a) A map of Israel shows the distribution of CWCN links in 2012. Different networks are depicted in different colors. Green: Cellcom (sampled once per day, magnitude resolution 0.1 dB); red: Cellcom (sampled once every 15 minutes, magnitude resolution 1 dB); blue: Cellcom (sampled once every 15 minutes, magnitude resolution 0.1 dB); black: Pelephone (sampled once per minute, magnitude resolution 1 dB); and orange: Orange (sampled once per day, magnitude resolution 0.1 dB). (b) The monitoring system devised by Giuli et al. [7].

backhaul links was then proven to be feasible by Zinevich et al. [11]. The logs containing the RSLs were processed and converted, using the $A-R$ law, to depict rain rates along a link. These values were then spatially interpolated to reconstruct a rain map.

COMMERCIAL WIRELESS NETWORK MEASUREMENT SYSTEMS

Attempting to reconstruct rain maps using the data that were obtained from the CWCNs requires an understanding of the processing that each RSL value undergoes. Such processing clearly depends on the equipment that is used in the cellular network. Here, we describe a representative subset of the processes to which each RSL sample is subject.

As digitally stored data dictate, the RSL value undergoes quantization. The RSL values are often saved after being quantized to a resolution of 1 dB, but a quantization of 0.1 dB may also be commonly found. The effects of the atmosphere and weather on the performance of a millimeter-wave communication link have been analyzed by Frey [12]. It has been found that the attenuation due to heavy rain at frequencies below 1 GHz is negligible. In fact, the rain-induced attenuations are in the order of the quantization and hence may not be measured properly using CWCNs. However, at frequencies above 15 GHz, the attenuation as a function of the rain rate is large enough to be measured. At frequencies of around 20 GHz, the attenuations go beyond the quantization magnitude, which enables a proper measurement of various rain rates using the CWCNs.

The backhaul operating frequencies of cellular networks vary depending on the communication technology. These are usually in the range of 20 GHz for longer-range links and may reach up to 40 GHz for short links where two antennas are closer together. This means that we may indeed use the logged backhaul RSL samples to measure rain rates.

The sampling rates of the RSL greatly vary from once per minute to a mere once per day. Sampling the minimal and maximal RSL data in a 15-minute interval is also common. In such cases, one must take into consideration that sampling the minimal and maximal values is a nonlinear process, making the reconstruction algorithms much more complex.

A central difference between the traditional monitoring methods and the CWCN sampling process is the fact that RSL attenuations are a product of an integration of the rainfall along a linear path. This is a result of the fact that the communication signal is transmitted using a highly directional antenna. As a result of the directivity of the antenna, raindrops cause interference to the signal when they enter its path, which may be modeled as propagating along a line. Rain along a line on which the signal propagates is the cause of the attenuation of the RSL. However, there is no reason to consider the rain rate constant along such a line. The sampled RSL is the integration of all rain-induced attenuations along a line connecting two antennas. Variations along the line on which a projection of the rain field has been applied may or may not be restored. Treatment of this issue is discussed in the

following sections. In this sense, it is clear that short links are preferred. Shorter links' RSL attenuations, on the other hand, are closer to the quantization magnitudes.

Attenuation along a link is naturally also caused by the propagation of the MW signal in space. To observe rain, we need to differentiate between any attenuation caused by nonprecipitation and the attenuation caused by precipitation. To do so, we need to measure the RSL data during times when no rain was present. However, these values also tend to vary. Wind, which moves the antenna, scintillation effects, temperature drifts, and other atmospheric conditions are the causes of these variations. The calibration of the attenuation level during times when no rain is present is usually named zero level or baseline determination and involves setting or choosing an RSL level that includes attenuation from all sources except the rain-induced ones.

Figure 2 shows the RSL data from two links that are located in Ramle, Israel. These RSL samples were taken during the same time, in two links that are roughly 2 km apart. One may easily notice that the signal strengths, given in decibel milliwatts, are different. While one link exhibits signal levels that decrease about 40 dB, from about –40 to –80 dBm, the second link introduces a dynamic range of only about 6 dB, from –35 to –41 dBm. The signal drop occurs at slightly different times but depicts the same rain event. These differences are mostly due to the different link lengths or link frequencies. We stress that the vast difference in RSLs is also due to the difference in link lengths and does not necessarily imply vastly different rain rates.

One may also notice the ringing effects in the short link [Figure 2(b)] that are caused by the 1-dB link quantization incurring quantization noise. To detect the rain event, a zero level of –40 dBm may be chosen in the left link's RSL. The added attenuation may be attributed to the rain event and may be converted to rain rate. However, in the right link's RSL, a zero level of between –35 and –37 dBm may be a good choice. The added attenuation is in the range of 5 dB. So the zero-level choice range is in the order of the added attenuation due to rain. Indeed, 1–2 dB of error in the calibration of the zero level seem negligible. However, a common link, operating at 15 GHz (which implies that the power law coefficients to use are $a = 35.7 \times 10^{-3}$, $b = 1.12$) will cause an error of 4.65 mm/h for a 5-km link length as

$$R = (A/La)^{1/b} = (1[\text{dB}]/(5[\text{km}] \cdot 35.7 \times 10^{-3}))^{1/1.12}$$
$$\cong 4.65[\text{mm/h}].$$

Hence, the calibration of the zero level is crucial for proper rain-rate measurement. This figure also depicts the temporal
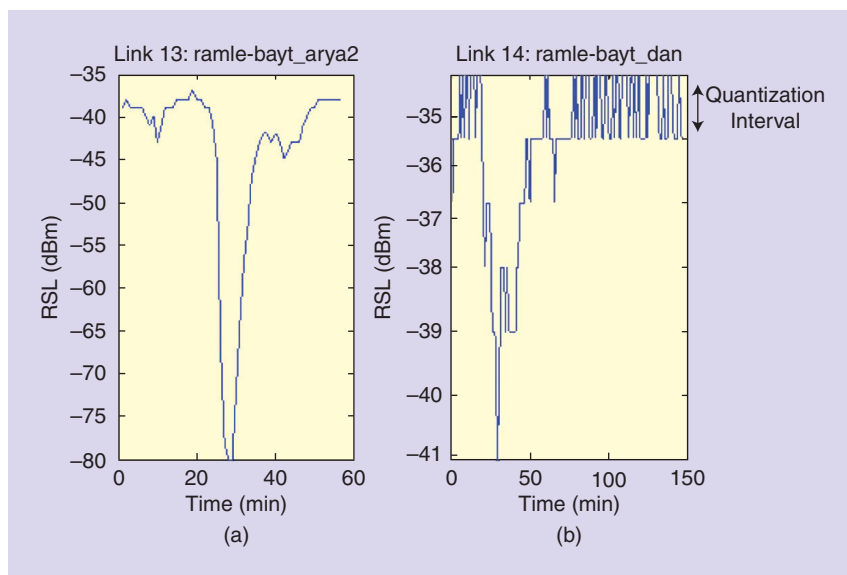
[FIG2] The RSL data of two nearby links: (a) link 13 and (b) link 14. Both links are located in the city of Ramle in Israel. The RSL data depict a rain event. Link 13 is 18.36 km, whereas link 14 is 4.55 km long. Neither link uses automatic power control.

and spatial dynamicity of the rain. After having evolved along a 2-km path, from the first link to the second, the second link measured an event that is distinctly different (the difference in link lengths is too small to be attributed solely to the difference in RSL measurements).

Fluctuations in the zero level are mainly attributed to variations in the water vapor density, ducting, and atmospheric scintillations [13]. Additional sources include changes in temperature that cause MWs to bend their propagation direction as a result of the change in the air's refractive index [12]. Winds that cause antenna deflections also result in RSL fluctuations [16].

OPPORTUNISTIC WIRELESS SENSOR NETWORK

During the past decade or so, advances in wireless communications have allowed the development of low-power, low-cost sensors built for the task of general-purpose sensing. Such sensors are found today in various applications, ranging from soil analysis [17] to the monitoring of sensitive wildlife and habitats [18], rainfall monitoring [19], and many more.

The desire to monitor phenomena for a long period of time, combined with the fact that, in many cases, the exact moment when the monitored phenomenon occurs is unknown, poses challenges in the energy budget of each sensor. Cases where the sensors cannot be replaced or treated often call for smart power management schemes.

In many wireless sensor networks, many nodes are deployed over a large area. To reduce the power consumption caused by the need to transmit the measurement results back to a base station (which may be located far from the sensor), the sensors can communicate with each other and deliver messages back and forth from other sensors. By doing so, real-time data over a wide area can be sampled. This implies that a common denominator in wireless-sensor-network-based

applications is the need to deploy a large amount of sensors. This increases the need for a cost optimization.

A central deficiency in such networks is caused by the need to monitor large areas. Because of the large number of deployed sensors, which are densely deployed in the monitored area, a substantial data redundancy among the sampled data will be present. And again, transmitting these data to a base station consumes energy and bandwidth. It is therefore necessary to develop efficient ways in which nodes can collaborate to send the relevant data only once.

According to [20], there are two main issues that rightly attract attention and differentiate sensor networks from the networks we know and understand: the limited power consumption and the potential to deploy networks with a large number of sensor nodes.

Corke et al. [21] studied a number of technological challenges that wireless sensor networks have presented in the past years and concluded that, in retrospect, the factors that have been found most critical to the applications' success are the optimization of power consumption and the need to efficiently cover a wide area of interest.

In fact, algorithms for optimization of the sensor topology, in an attempt to minimize redundant data and/or power consumption, are an active research area (see, e.g., [22]–[25]). A thorough survey discussing wireless sensor networks may be found in [26].

The newly suggested CWCN measurement system may be considered a wireless sensor network. It consists of a multitude of sensors that may be used for the purpose of monitoring the environment. However, a fundamental difference distinguishes it from other wireless sensor networks. The sensors are already deployed, have zero cost, and are fed by an infinite power source. This is the reason we call such a network an *opportunistic wireless sensor network* (*OWSN*).

Many inherent characteristics differentiate the OWSN and the typical wireless sensor network. In a typical sensor network, the sensors are optimized for the task of monitoring a specific phenomenon. In the OWSN, no such optimization can take place. The sensors are optimized for communication quality of service. Their sensitivity, for example, is far from optimized for the monitoring task. For example, if we were to attempt to monitor a slight drizzle by measurements taken from a sensor operating at 26 GHz, we would need a measurement resolution or quantization of ~0.02 dB (see [12]). Common sensors in OWSNs have a sensitivity/quantization of 0.1 or 1 dB. Such limitations may be alleviated by considering the amount of data available. It is reasonable to believe that the large quantization, for example, may be mitigated by averaging the measurements of a large amount of sensors. Some of the papers we describe in the following sections have accepted the sensors with their many limitations and turned to statistical signal processing tools to cope with them.

Another central limitation from which the OWSN suffers is the highly irregular manner by which the sensors are spread across land. This is depicted in Figure 1. This presence of areas with insufficient coverage gives rise to the problem of data assimilation. The sampling of insufficiently covered areas may be achieved by assimilating rain gauge data and/or weather radar or even data from different commercial network service providers.

Another factor that makes the monitoring task difficult is the fact that the communication systems employ an automatic power control and adjust the transmission power according to the measured signal power. This automatic tuning must be taken into account when attempting to infer the proper values of the monitored phenomena. Table 1 summarizes the central differences between the newly suggested OWSN and typical wireless sensor networks.

SOURCES OF ERRORS

A systematic source of error in the monitoring and observation of rain with CWCN is due to the approximation that yielded (1). Furthermore, the calibration of the a, b coefficients for the $A - R$ relation must be carefully applied for gaining a proper rain-rate measurement from the RSL data. If we were to properly measure the rain-induced attenuation, we would apply a relation that integrates the rain along the path that connects two links, rather than assuming that the rain is constant along such a line.

The question of how we are to consider the rain rate that we measured using an RSL reading arises. The returned rain rate is usually treated as a path-averaged rain rate along this line. However, to the best of our knowledge, no reconstruction algorithm currently suggests a method for an exact reconstruction of the rain rates along the path of integration. Some algorithms do divide a line into several points [13], [14], but no algorithm fully reproduces the rain rates along the line in a continuous manner. The analysis of the ability to reconstruct rain maps by Sendik et al. [15] hints at methods to do so. Consequently, an algorithm that postulates that the rain rate along the path is equal to the path-averaged rain rate is probably erroneous. This suggests that the longer the path along which the rain rate is considered to be constant, the larger the errors.

Longer link distances are usually found in more rural areas, where a smaller population density makes use of the cellular network. This means that, usually, when links are long, they are also less dense and render monitoring algorithms prone to errors. Figure 1 shows that in the northern, southern, and eastern parts of Israel, where population densities are lower, longer links are more common, and their density is noticeably lower, whereas in the central part of Israel, where the population density is high, the links are short and quite dense.

[TABLE 1] OPPORTUNISTIC VERSUS TYPICAL WIRELESS SENSOR NETWORKS.		
CHARACTERISTIC	**OPPORTUNISTIC**	**TYPICAL**
POWER CONSUMPTION	NO LIMITATION	A CENTRAL LIMITATION
AMOUNT OF SENSORS	THOUSANDS	TENS TO HUNDREDS
SENSOR COST	ZERO COST	OPTIMIZED FOR COST
DEPLOYMENT COSTS	ZERO COST	OPTIMIZED FOR COST
MEASUREMENT SENSITIVITY	INSENSITIVE	HIGHLY SENSITIVE
GEOMETRIC DISTRIBUTION OF SENSORS	NO CONTROL	OPTIMIZED FOR COVERAGE
*Blue indicates good and red indicates poor characteristics.		

As mentioned previously, the temporal sample rates of the RSL vary greatly depending on the type of equipment used by the network service providers. Sample rates as low as once per day are incontrovertibly inadequate for reproducing rain maps. Perhaps the 15-min resolution sample rate is sufficient for some applications. This is yet to be determined, as the number of reconstruction algorithms that account for temporal effects is still small [27].

The highly dynamic spatial and temporal natures of rain couple between sources of error, which are caused by the link averaging because of its length and the sample rates. By simulating CWCN RSL data from path-averaged radar samples, both temporal and spatial errors were analyzed by Leijnse et al. [28]. Applying a sample rate of 15 min, they have shown that errors increase with link length, as the dynamic spatial nature of rain causes an increase in the rain-rate variations along the link. They also showed that different sampling strategies have a crucial role in controlling sources of error. For example, they proved the inferiority of a sampling scheme that simply returns an instantaneous RSL, when compared to a time-averaged RSL. An analysis of the errors due to the spatial variability of rainfall was applied by Berne et al. [29]. By using a stochastic simulator of the DSD, they analyzed the influence of the link's frequency, length, and DSD spatial variability on the rainfall estimation. They showed that the error due to the usage of the power law to connect the RSL and the rain rate is negligible for frequencies between 10 and 50 GHz for links longer than 15 km. However, in urban areas, links that span only several kilometers may induce errors of up to 4%. Zinevich et al. [30] showed that the most dominant source of error (assuming an effect called the wet-antenna effect, which we discuss in the following paragraph, is corrected for) is the spatial nature of the rain, surpassing the errors that are caused by quantization, zero-level uncertainty, DSD variability along the link, and others. Alleviating this source of error requires using a large amount of densely distributed sensors, a requirement that is easily achieved using the CWCN-based monitoring system.

The proper calibration of the zero level is also a crucial factor for an accurate rain measurement. Events with low rain rates induce only a minor additive attenuation and, hence, require exact zero-level calibration to enable their detection. Longer links have more rain along their path, causing larger signal attenuations. This eases the detection of low rain rates.

Another widely discussed source of error is the wet-antenna effect. Humidity sources cause moisture to accumulate on the antenna radome and cause an added attenuation. Applying a model originated by Kharadly and Ross [31], Minda and Nakamura [32] have suggested an exponential model relating the actual rain attenuation to the total attenuation induced by both rain and the wet-antenna effect. Their equation suggests that, for a constant amount of accumulated humidity on an antenna, a constant value of added attenuation is caused. Hence, time-averaging the RSL values in the process of zero-level calibration will not cancel this effect. In [28], it was shown that the wet-antenna effect is most probably the greatest source of error for short links. The added attenuation is in the order of magnitude of common rain events and completely biases the rain-rate measurement. Schleiss et al. [33] have shown that the wet antenna, which affects the CWCN links, increases in an exponential manner during rain and decreases exponentially back toward zero once the rain stops.

The RSL values are logged after being quantized. The received strength levels given within a resolution of 0.1 dB are surely precise enough to measure rain rates. Precipitation other than rain, such as fog or dew, however, generates attenuations that are significantly less than those caused by rain. Hence, the RSL's quantization is a source of error that must be considered before attempting to observe such phenomena.

Often, nonlinear processing is applied to the RSL samples before logging. An RSL that is sampled once every 15 min commonly undergoes a min/max thresholding. In other words, only the minimal and/or maximal value of the RSL is saved every 15 min. It is indubitable that such nonlinear processing applied on a signal may incur reproduction errors. For short link lengths or links using low frequencies, the natural fluctuations of the zero-level attenuation have the same order of magnitude of a quantization interval of 1 dB. Because of the nonlinear processing in addition to the quantization, the error in the baseline estimations may affect an entire rain event, which may introduce a bias in the estimation of rainfall.

To summarize, following the analysis applied by Zinevich et al. [13], the most dominant source of error is the spatial rain variability, which causes errors if the CWCN is not distributed in a sufficiently dense manner. Hence, this may be relatively easily resolved within areas with a high density of links that have a wide range of lengths. The second source of error in magnitude is the zero-level choice, which must be carefully calibrated. After properly calibrating the zero level, the DSD and wet-antenna effect are the most dominant sources of error, followed by the quantization of the RSL values. Table 2 presents a comparison between the properties of the CWCN and traditional precipitation-monitoring systems.

SIGNAL PROCESSING

CALIBRATION

As previously mentioned, estimating rain rates from the RSL data requires calibrating the baseline or zero level. Various techniques for such a calibration may be found in the literature [34]–[39]. Perhaps one of the most paramount advantages of the CWCN approach to precipitation monitoring is the presence of a multitude of data. Using a network of sensors (many links) that sample the same rain event may be of help when attempting to determine the baseline. Methods that make use of more than one RSL time series for properly calibrating the baseline are presented in [14] and [34]–[37].

The first to show the advantage of using two links with close frequencies were Rahimi et al. [34]. They suggested the use of dual-frequency MW links for measuring path-averaged rainfall. They presented a baseline determination method, which leaned on the assumption that rainy periods are short. These methods included a self-updating baseline attenuation level for each frequency during the dry periods. The latest dry period was used as a baseline level for the upcoming rainy period, which may then be refined using the subsequent dry period. For this, they need to properly detect wet/dry periods. Hypothesising that the correlation between RSL

[TABLE 2] THE CHARACTERISTICS OF PRECIPITATION MEASUREMENT SYSTEMS.

	RAIN GAUGE	WEATHER RADAR	SATELLITE	CWCN
PRICE (PER UNIT)	LOW	MID	HIGH	VERY LOW
DEPLOYMENT COMPLEXITY	HIGH	MID	HIGH	VERY LOW
CALIBRATION COMPLEXITY	LOW	MID	MID	MID
DATA PROCESSING COMPLEXITY	LOW	MID	HIGH	MID
SPATIAL RESOLUTION	LOW	MID	MID	HIGH
TEMPORAL RESOLUTION	LOW	MID	MID	HIGH
SHORT/LOCAL MONITORING SUITABILITY	LOW	MID	MID	HIGH
LONG/GLOBAL MONITORING SUITABILITY	HIGH	MID	HIGH	HIGH

time series of the two frequencies during wet periods is higher, they found that

- in dry events, the median of the correlations is under 0.25
- in wet events, the median of the correlations is above 0.79.

They then calibrated the baseline value, choosing the attenuation measured just before a wet event, and they refined it using the baseline that preceded the wet event. Overeem et al. [35] proposed a method for zero-level calibration in the case of min/max RSL data. The RSLs of 57 commercial MW links around the city of Rotterdam, The Netherlands, were sampled during 15-min intervals, and the minimal and maximal values were logged at 0.1-dB resolution. Analyzing links shorter than 10 km, they defined two terms:

- the difference between the instantaneous minimum RSL and the maximum of the minimum RSL during a 24-h period: $\Delta P = P_{\min} - \max_{24\ h}\{P_{\min}\}$
- the link length normalized difference: $\Delta P_L = \Delta P/L$.

They identified a transition to a wet period by requiring that the medians of ΔP and ΔP_L be under predetermined values. Such a transition into a wet period was extended as long as

$$\max_{24\ h}\{P_{\min}\} - P_{\min} > 2\ \text{dB}.$$

Hadar [36] suggested employing hidden Markov models (HMMs) to identify dry/wet periods. The RSL measurements were the HMM observations and the hidden state was either wet or dry. Having detected dry and wet periods, the baseline was set to the value just before a transition from a dry to a wet period. The CWCN RSL data, after subtracting the zero-level attenuation for proper rain-rate measurement, were correlated to rain gauge data and were found to correlate well, yielding correlation values at about 0.7. Yet another HMM-based approach for inferring dry and rainy periods from telecommunication MW link signals was suggested by Wang et al. [40].

Methods for baseline determination that do not incorporate the multitude of data inherent in CWCNs are also found in the literature [38], [39]. Schleiss and Berne's [38] method of differentiating between dry and rainy periods comprised a calculation of the standard deviation of the RSL data in a predetermined window of 15–35 min. Chwala et al. [39] suggested a spectral approach that involved applying a short-time Fourier transform to the RSL signal and considering its power spectrum. Dividing the spectrum into a low region and a high region, the power in each region was calculated. If the difference between powers in the two regions exceeded a preset threshold, the event was considered a wet one. This is motivated by the hypothesis that rain events impel high-frequency RSL samples.

Kaufmann and Rieckermann [41] have discussed three different methods for baseline determination. These included 1) a moving window algorithm, 2) a statistical classification algorithm using random forests, and 3) an algorithm based on a Gaussian factor graph. The first method, which included a moving window algorithm, is, in essence, a modification of the algorithm that was previously suggested by Schleiss and Berne [38]. The second method, which included random forests, required defining a set of attributes or properties of the RSL data from which a classification into a wet or dry event may be applied. These attributes are then thresholded while entering a tree of classification decisions. A tree leaf yields a final classification. The third method, which they applied, was based on the Gaussian factor graph. This approach involves modeling the rain process in the state space in which the state space vector was chosen to include the RSL value and its slope. By recursively relating between past and present RSL observations, an RSL sample may be effectively denoised and then reconstructed. This in turn enables classifying the event as either wet or dry. A dry classification means that the current reconstructed sample is part of a baseline. Important assumptions for their approach are that the data belonging to the baseline are locally smooth and periodic.

Holt et al. [42] have determined an RSL baseline by employing the assumption that the RSL data from two frequencies are very highly correlated during rain events. They therefore classified events as dry in cases where correlations were below 0.8 and where there was no record of any rainfall at added rain gauges.

Another effect that raises the need for calibration is the wet-antenna effect. Zinevich et al. [13] calibrated the wet-antenna coefficients, assuming its independence in frequency, by using rain gauge data as ground truth for rain rate. These coefficients were found optimal for accounting for wet-antenna effects during rain (due to accumulation of raindrops on the antenna radome during the presence of a rain event). However, these coefficients are inappropriate when accounting for wet-antenna attenuation, which is caused due to nightly dew or any other source of accumulated drops. In general, the correction and calibration of the wet antenna effect is understudied and requires a more profound study, especially after acknowledging the findings in [28].

Other than the zero level and the wet-antenna effect, the power law coefficients must also be calibrated. However, most reconstruction and/or estimation algorithms make use of values similar to those suggested by [10]. Common algorithms do not correct for temperature drifts or DSD variations.

DETECTION

Once a CWCN system has been calibrated for proper measurement, we may use it to detect various types of precipitation. For example, the ability to provide essential rainfall information from regions prone to flash floods was exemplified by David et al. [43]. Preliminary results concerning fog monitoring using commercial MW systems have been shown by David et al. [44], applying the Rayleigh approximation to relate fog to attenuation per kilometer.

The detection of vegetation diurnal cycles by using a custom-built CWCN was demonstrated by Hunt et al. [45]. Setting out to measure vegetation characteristics, they deployed a network of seven rain gauges in a cornfield north of Ames, Iowa. To avoid the manual collection of data from the gauges, they transmitted the rain gauge data to a tower. Coincidentally, they found that the signal strengths reveal cycles. Investigating these cycles, they found that the RSLs indicate whether vegetation is present in the signal propagation path. Differentiating between periods where the vegetation was harvested and periods before the harvest, they found the RSLs to have considerable differences and used these differences to prove that the signal strengths may be used for vegetation monitoring. Moreover, they showed that the signal strength is inversely proportional to the vegetation water content. Harel et al. [46] applied an extended multifamily likelihood ratio test for precipitation detection, discriminating between wet and dry periods.

ESTIMATION

Estimating rainfall rates from RSL samples has been treated both by approaches that make use of the already deployed CWCNs [6], [37], [47]–[51] and by approaches that deploy custom-built equipment [34], [39], [48].

In essence, the problem of transforming RSL values to rain rates is simple. It most practically involves setting a baseline, choosing power law-coefficients, and applying the inverse of the power law given in (1). The main drawback of this method is that it implicitly states that the rain rate along the link line is constant.

However, in practice, one must consider other phenomena that may sabotage such straightforward attempts. The wet-antenna effect, outlier samples, mismatches in the power law coefficients and many more real-life processes may affect the RSL data and result in incorrect rain rates.

Leijnse et al. [6], being one of the first to present actual rain-rate estimation from CWCN RSL data, have recognized a systematic overestimation, which they reasoned is partly due to the uncertainty in the baseline signal level settings, but they stated that it is more likely the result of extra attenuation caused by the wet antennas, which can cause several dBs of additional attenuation [31].

Having acknowledged the baseline determination issue, Rahimi et al. [34] applied two different methods for zero-level setting on custom MW equipment that measured signal attenuations. To properly asses the feasibility of their custom-built equipment, they searched for a method to compare MW path-averaged samples to rain gauge point samples. To do so, the rain-gauge data were converted into path-averaged data by allocating each portion of the link to its nearest gauge. However, one must comprehend that such a conversion is required for the sole purpose of comparison and observation quality assessments. For a fully operational system that observes rain rates from CWCN data, no such conversion is required.

Kuntsmann et al. [48] recognized the fact that CWCN-based precipitation observation systems may be of great assistance in regions with either a course station network density or high spatial precipitation variability, and stated that the water resource management community may be greatly aided by CWCN-based monitoring systems. They applied a CWCN-based precipitation-monitoring system in an orographically complex terrain, the prealpine region of southern Germany, where precipitation fields derived using radar data are erroneous. This is due to the inability of the radar signal to track the terrain slopes. They set out to build a cell phone provider-based system reinforced by hydrological and meteorological radar and rain gauge data from an observation site. For the purpose of feasibility studies, they built a polarimetric transmission device, set to investigate the interaction of MWs with precipitation.

Interested, too, in the alpine and prealpine region of southern Germany, Chwala et al. [39] used custom MW-based equipment and applied their novel baseline determination technique for precipitation observation. When comparing to rain gauges, they succeeded in acknowledging the fact that an indication of dry periods by rain gauges does not necessarily mean that there was no rain along the link at all; the link RSL data represent a path-integrated rain rate, whereas gauges are point samples.

Rayitsfeld et al. [37] compared two methodologies for long-term rainfall monitoring by CWCNs. Their first methodology used simple RSL data from a single link, applied the power law, and compared the outcome results to data from the closest rain gauge. The second methodology followed Goldstein et al. [14] and used a modified inverse distance weighted interpolation to calculate rainfall at the rain gauge point based on the RSL values from all of the nearby links. In general, the results indicated that the two methods improve as the density of the links increases, which is most probably one of the cardinal advantages of CWCN-based monitoring systems, a multitude of links or sensors. However, strictly speaking, such methods are not considered estimation techniques but rather reconstruction methods as they involve the generation of new data (such as rain rates in locations where links do not exist). The method described by Goldstein et al. is discussed briefly in the following sections.

Ostrometzky [52] established a method for robust precipitation estimation, regardless of the specific water phase (liquid, solid, or a mixture of both). Ostrometzky, having recognized the function of attenuation versus snow rate A_{Snow}, which was given by Frey [12], suggested a simple additive attenuation model,

$$A[\mathrm{dB/km}] = A_{Rain} + A_{Snow} + \gamma(A_{Rain}, A_{Snow}),$$

where A_{Rain} is given by the power law in (1), and $\gamma(A_{Rain}, A_{Snow})$ is a sleet interaction term, which causes sleet-induced attenuation.

Then, by exploiting the presence of a multitude of data, Ostrometzky applied a least-square process to estimate the rain, snow, and sleet rates.

Luckily, part of the commercial MW links have a quantization error of $0.1[dB/km]$. David et al. [43], [53] realized that during typical conditions the attenuation caused by the water vapor is $\sim 0.2[dB/km]$ and exploited this for estimating water vapor. However, the technique suggested by David et al. is restricted to weather conditions that exclude rain, fog, or clouds along the propagation path, and the determination of the RSL zero level is done using side information. In other cases, a classification or separation phase must be invoked before the estimation attempt.

CLASSIFICATION AND SEPARATION

Once attenuation is introduced by precipitation, the use of the power law for converting RSL into rain rate is applicable. However, how is one to know that the observed precipitation is, indeed, rain? The research of MW attenuation by precipitation has not been limited to rainfall. Cherkassky et al. [54] have proposed a detection/classification system capable of detecting wet periods, with the ability to classify the precipitation type as rain or sleet (a mixture of rain and snow), given an attenuation signal from spatially distributed CWCN links. They divided the classification process into two stages. In the first, events are classified into wet or dry events. Then, in the second stage, wet events are further classified into sleet or rain events. Cherkassky et al. have used the RSL signal features (such as fade duration, fade magnitude, and fade slope) for the purpose of classification. Following the assumption that sleet and rain events may be distinguished by observing the fade dynamics, a feature vector for classification was chosen.

ASSIMILATION

Often, various types of RSL samples are present. CWCNS, which are constructed using various manufacturers, may cause the multitude of RSL samples to consist of differently sampled data. This may also be caused when one is attempting to reconstruct rain maps while using data from various service providers at once. In Figure 1, a map of the MW links in Israel is depicted. The green links are sampled once per day with 0.1-dB resolution, red links are sampled once per 15 min with 1-dB resolution, blue links are sampled once per 15 min with 0.1-dB resolution, and black links are sampled once per minute with 1-dB resolution. Thus, if we are to reconstruct a rainfall map out of the multitude of links, we must consider the problem of assimilating the various types of data into one map.

To the best of our knowledge, the problem of assimilating the various types of data has not been treated in the scope of rain monitoring using CWCNs. CWCN-based observation systems will probably not be able to monitor the oceanic regions of the earth. This is because cellular antennas are not deployed over international regions in general and over oceans, in particular. This fact may be a cause for the need to assimilate between satellite and CWCN data for a global outlook on weather. Again, to the best of our knowledge, no research work has been applied on the problem of assimilating CWCN data from traditional observation tools.

RECONSTRUCTION

Algorithms that attempt to reconstruct rain maps from RSL data are most easily found in the literature. Algorithms that do not account for any temporal evolution of convective clouds but rather display a converted value of the instantaneous RSL to rain rates are given in [11] and [14].

In [11], Zinevich et al. have proposed a nonlinear tomographic model that treats the problem of the variability of the cell sizes (a cell is defined as the area enclosed between CWCN links), accounting for the irregularity of the network topology, observation quantization, and nonlinearity of the power-law equation for different links.

Their algorithm begins by dividing an area covered with links into cells. Conventional tomographic algorithms use rectangular grids, which do not fit in this case, as the spatial distribution of the links is highly irregular. Such an algorithm will benefit from the fact that a relatively constant number of links will appear in each final cell instead of having cells with many links in urban areas and cells with few links in rural areas.

Figure 3(a) depicts the radar rain map, Figure 3(b) the CWCN-based map, and Figure 3(c) the division of the area into cells by the algorithm above. A general consent between the radar map and the CWCN-based map may easily be noticed.

Another algorithm for rain-map reconstruction was suggested by Goldstein et al. [14]. The proposed algorithm consisted of preprocessing the links' data, followed by a weighted least-squares algorithm to extract the rain level at any given point in space. In Goldstein's approach, similar to that of Zinevich, each link was divided into K intervals in order not to impose the constancy along the link. Each rainfall value is then reconstructed by using more than one point in space, taking into consideration neighboring links. A weighting of the original rain rates, which are attained from the RSL values, was also applied. The weight was chosen as an inverse of the point's distance from an actual link. A functional was then iteratively minimized for the sake of rain-map reconstruction.

More approaches to the problem of reconstructing a rainfall map from MW link measurements have been suggested by Overeem et al. [47], [49] and Watson and Hodges [55]. Overeem et al. suggested a method for reconstructing countrywide rainfall maps from CWCNs and applied it on minimum and maximum RSL samples with a temporal resolution of 15 min. Having adjusted the RSL levels by rain gauges and path-averaged radar, rainfall intensities were derived.

Watson and Hodges [55] formalized the reconstruction problem as a problem of finding an orthogonal basis of functions, which spans a rain-field function. They were then left with extracting the coefficients, which are the projections of the rain map onto the basis functions. These coefficients then enabled a proper reconstruction of the rain map as a linear combination of the basis of functions, which spans the rainfall map solution by applying a least-squares technique.

None of the previously mentioned reconstruction algorithms exploit the temporal nature of rain fields. An algorithm that depicts the temporal evolution of the rain fields was suggested
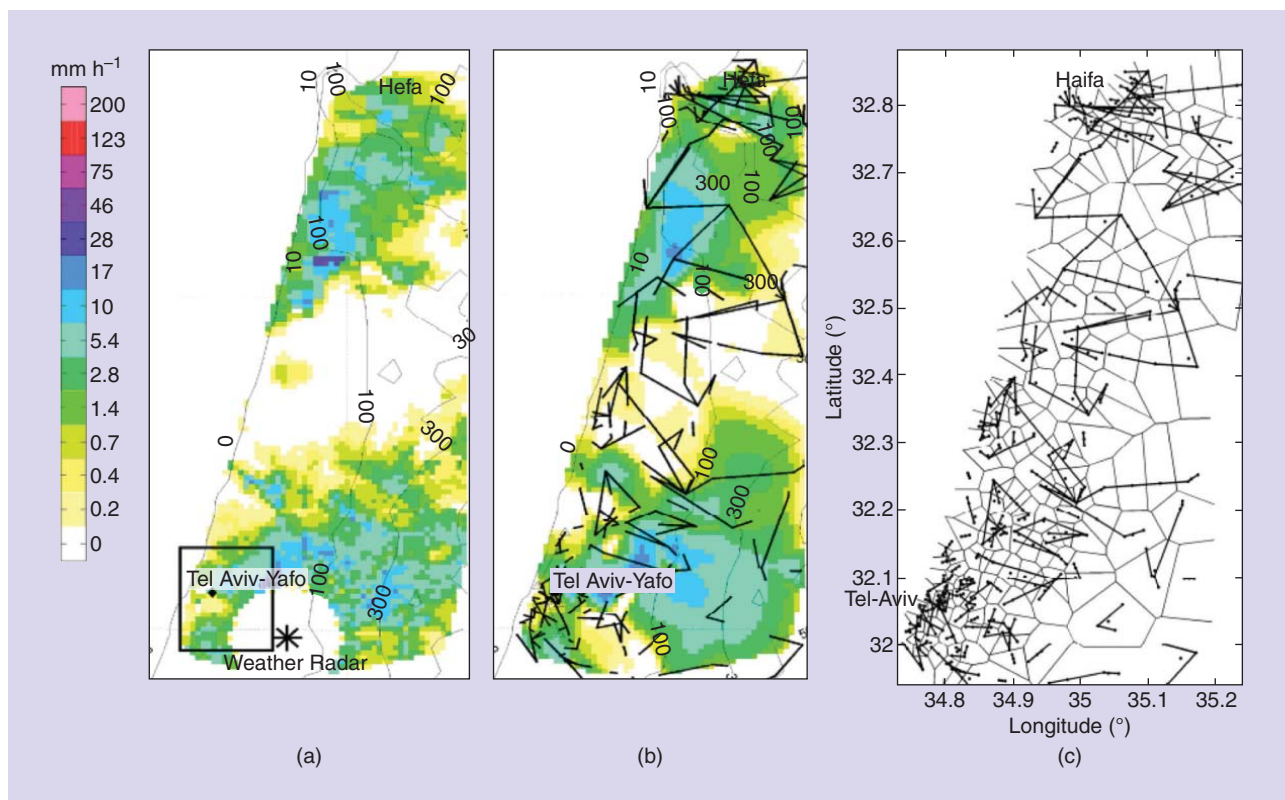
[FIG3] A reconstruction of rain maps: (a) a radar-based map, (b) a CWCN-based map, and (c) the location of the links in the CWCN-based map [11].

by Zinevich et al. [27]. Assuming a translational rain field evolution model, under the assumption that the rainfall advection is driven by wind, the extended Kalman filter was used to determine the wind velocity and direction at the storm steering level using the RSL data. By doing so, Zinevich explored the concept of recovering the rainstorm dynamics from CWCN links.

Another algorithm specifically developed for CWCNs in urban areas, where a large number of MW connections is typically found, was presented by Cuccoli et al. [56].

Sendik and Messer [15] addressed the problem of the ability to reliably reconstruct a two-dimensional function (e.g., a rain map) being sampled by projections along lines, without any restrictions on the line types. They then applied their analysis to the problem of reconstructing rain maps from CWCN links. Their solution to the question regarding the ability to reconstruct a two-dimensional function, which is sampled by an arbitrarily set of lines involved employing a series of three separate stages, which consist of first solving a problem of sampling with a regular grid but with arbitrary types of lines. In the second and third stages, they portrayed the problem of a nonregular grid as one with missing samples. Essentially, these three stages enabled the consideration of the CWCN sampling scheme as a case of regular sampling with missing samples, which have been sampled by a linear functional, which is the mathematical representation of the line along which the projection had occurred.

Applying these three stages yielded an answer stating whether the set of links can be used for reconstructing rain maps without errors. The process above also enables the derivation of the maximal frequency, which can be sampled without causing any aliasing errors. This work, which made use of the Papoulis generalized sampling expansion stating the exact reconstruction kernel, may, perhaps, be used for future algorithms that attempt to reconstruct rain maps.

OPEN CHALLENGES

SENSOR-BASED CHALLENGES
Previous sections dealt with a wide range of signal processing challenges to which the usage of RSL data gave rise. However, many of the algorithms and/or techniques described above still do not truly use the vast amounts of RSL data that are present in urban areas.

The raison d'etre of the CWCN, in our opinion, is the availability of a large amount of data that must be exploited for generating robust and exact estimates. Linking between sensor networks and CWCN-based precipitation-monitoring systems is a direction that should be fully examined and exploited. It is generally acknowledged that the advantages of using sensor network techniques include the ability to cope with failures of sensors, robustness to outliers, and the ability to monitor a wide variety of phenomena through the application of statistical signal processing methods (based on the multitude of data). This avenue is yet to be examined in the scope of CWCN-based monitoring systems.

A CWCN is composed of low-cost, spatially distributed autonomous sensors. Hence, it is, in essence, a wireless sensor network that enables using tools and results from the wireless sensor network community. Using the CWCN sensor network as an opportunistic one and, in particular, its diversity of measurements, should enable applications that a single senor simply cannot provide.

A major challenge is the study of the dependencies between an actual situated network of sensors topology and the monitoring accuracy. It is intuitive that the denser the network of sensors, the better the reconstruction accuracy. Hence, a central question is regarding the sufficient sampling set for properly reconstructing environmental phenomena sampled by an arbitrary geometry of a CWCN. An answer to this question enables the ability to both analyze existing CWCNs but also the ability to synthesize CWCNs. That is, in cases where the answer regarding the sufficiency of a CWCN is found to be insufficient, one may want to consider the locations to add sensors to render the CWCN sufficient for a proper reconstruction. For example, a rain gauge may be added in locations where the CWCN density is low and found to yield an insufficient sampling scheme.

This result also gives rise to assimilation challenges. In a case where a rain gauge was added to amend an improper sampling scheme of a CWCN, generating a map of the observed precipitation requires assimilating the rain gauge data together with the RSL data. While the idea of assimilating the measurements of a single link, rain gauges, and radar has been studied [57], [58], integrating an entire CWCN with other meteorological measurements is still an open research question.

CWCNs are often composed of links with various time and power resolutions. The question of how to assimilate these various CWCN links to reconstruct a single environmental phenomenon is yet another open problem.

The methods described by David [43], [44], [53], [63] prove the feasibility of using CWCNs for purposes other than rain-rate estimation. However, the techniques suggested by David are restricted to weather conditions that exclude rain or clouds along the propagation path. In essence, we feel that the classification and separation problem is still an open problem that must be treated in the presence of various precipitation types, all measured at once.

The problem of data assimilation seems inevitable because the growing popularity of CWCN observation methods will require large amounts of RSL data. Whether we wish to assimilate RSL data sampled by different sampling schemes or to reconstruct global rain maps, which require assimilating between satellite and CWCN data, a rigorous treatment of the assimilation of CWCN data is necessary.

Also, as previously stated, for precise reconstruction of rain maps, algorithms cannot impose constant rain rates along the path of the link. All of the algorithms to date do not address this issue.

VECTOR SENSORS

In most of the papers mentioned here, only the magnitude of the CWCN data was used, in other words, the RSL. However, electromagnetic waves are also characterized by their phase, which may be altered because of propagation effects. The benefits from making use of the phase data in CWCN-based data are yet to be determined. Bringi et al. [59] examined the propagation effects in rainfall on radar samples at frequencies of 3, 5.5, and 10 GHz, simulating the difference of attenuations and difference of phase between horizontal and vertical polarizations. They found that a near-linear relation exists between attenuation and the differential propagation phase.

This motivates the incorporation of vector sensors. In other words, a CWCN that logs both RSL and phases for more than one polarization may be found foundational for a stable, error-proof, precipitation-monitoring system. It may also be found that differential phase data suffer less from quantization errors and thus enable a precise monitoring of phenomena other than rain. As David et al. [53] stated, the attenuation caused by the water vapor is $0.2\,[\text{dB/km}]$, giving rise to the desire for more precise measurements, which may be attained by incorporating phase and polarization data.

MONITORING PHENOMENA OTHER THAN RAIN

In the sections above, we have discussed mainly the monitoring of rain. The fact that most of the works treat the problem of rain observation is not coincidental but is related to the fact that out of the variety of precipitation types, rain yields the largest attenuation amplitudes. However, as shown above, David et al. [43], [44], [53], [63], [64] suggested CWCN-based methods to observe water vapor and, potentially, even fog, and they have the potential to enhance the ability to cope with flash floods. To date, no rigorous treatment of the problem of observing hail, snow, graupel, or dew has been completed.

Errors in monitoring rain in heavy storms are caused by the slight shifts of the antennas by the storm winds. Monitoring winds by measuring the attenuations induced by these slight shifts may, perhaps, be found practical.

Phenomena other than precipitation may also be monitored by CWCN-based systems. Ghobrial and Sharief [60] discussed the electrical properties of dust and derived expressions for attenuation and phase shifts for a medium with precipitating dust particles in terms of visibility and wavelength for vertical and horizontal polarizations. They found that the dust-induced attenuation is related to the width of the dust layer by the wavelength and the visibility. They concluded that dust storms resulting in visibilities of 10 m or less introduce considerable attenuation. This may be a window to the detection of dust storms using CWCN RSL data.

Andrews [61] measured the absorption of MWs by both carbon monoxide (CO) and nitrogen dioxide (N_2O) at a frequency of 9.75 GHz and found that the power absorption was found to increase with density in both CO and N_2O. As is widely known, motor vehicle emissions are composed of CO and nitrogen oxides (N_2O and nitrogen monoxide). This may advocate the ability to monitor air pollution densities by a CWCN.

Burning wood reacts with oxygen, producing carbon dioxide (CO_2) and water (H_2O), which are both released as gases in

the air. The attenuation induced by water is the fundamental effect on which rain monitoring is based. Can this reaction give rise to CWCN-based systems for monitoring forest burns? Time is yet to determine whether or not the avenues suggested above will indeed evolve into operational monitoring CWCN-based systems.

CONCLUSIONS

We have presented the physical basis that led to founding the new precipitation-monitoring approach, the CWCN-based system. An aggregation of techniques and algorithms for making use of the CWCN RSL data may be found. These include methods for detection of wet versus dry periods, estimating local rain rates, reconstructing countrywide rain maps, detecting flash floods, fog, and more. Inherent error sources such as the proper treatment of the wet-antenna effect still need to be addressed to build proper and precise precipitation-observation systems.

The problem of data assimilation seems inevitable and will soon require a thorough understanding and treatment, if the CWCN-based system for precipitation monitoring is to replace the traditional systems. Such treatment must include both the assimilation of various RSL data types and the assimilation of CWCN RSL data with traditional sampling systems. The assimilation of various RSL data types requires either interpolating or downsampling of samples that have been sampled differently. The problem of assimilating between samples that have been processed nonlinearly (minimum and maximum values) is a more complex problem.

CWCN-based observation systems currently lack the ability to monitor international regions such as oceans. This generates the need to know how to properly assimilate between traditional monitoring systems such as satellite data and the CWCN RSL data.

The most striking issue, in our opinion, is the fact that none of the papers currently in the literature recognized that the CWCN is, in essence, a sensor network. As a result, none of the methods described has truly made full use of the potential that is hidden in the multitude of available data. Applying stochastic signal processing algorithms may enable more precise, robust, and stable reconstruction algorithms.

Hints on the feasibility of the CWCN-based monitoring system to monitor phenomena other than precipitation were discussed. These included fire detection, pollution detection, and perhaps even dust-storm detection. All of these require only RSL measurements or, in other words, only amplitude samples. If the CWCN someday logs phase data, its links may then be treated as vector samplers, which may be found to enable a wider range of applications. Yet, all of these new avenues are challenging and still far from being implementable. To conclude, we believe the CWCN system is only in its beginning, depicting only a very small portion of its full potential.

ACKNOWLEDGMENTS

AUTHORS

Hagit Messer (messer@eng.tau.ac.il) received her Ph.D. degree in electrical engineering from Tel Aviv University (TAU), Israel, and, after a postdoctoral fellowship at Yale University, she joined the faculty of engineering at TAU in 1986, where she is a professor of electrical engineering. She has published numerous journal and conference papers and has been a supervisor to graduate students. She has been a member of technical committees of the IEEE Signal Processing Society since 1993. She is on the editorial boards of *IEEE Transactions on Signal Processing*, *IEEE Signal Processing Letters*, and *IEEE Journal of Selected Topics in Signal Processing* and on the overview editorial board of the IEEE Signal Processing Society journals. She is an expert in statistical signal processing with applications to source localization, communication, and environmental monitoring. She is a Fellow of the IEEE.

Omry Sendik (omrysendik@mail.tau.ac.il) received his B.Sc. degree (with honors) in electrical engineering and his B.A. degree (with honors) in physics, both from the Technion–Israel Institute of Technology, Haifa, Israel. He received his M.Sc. degree (with honors) in electrical engineering from Tel Aviv University. He received the Excellent Project Award of the Communications Laboratory, Technion–Israel Institute of Technology, for his work on cognitive radios. He was also awarded the Annual Intel Award for Excelling Science and Engineering M.Sc. and Ph.D. Students.

REFERENCES

[1] E. C. Barrett and M. J. Beaumont, "Satellite rainfall monitoring for agrometeorology: operational problems, practices and prospects," *EARSeL Adv. Remote Sensing*, vol. 2, no. 2, pp. 66–72, 1993.

[2] A. K. Biswas, "The automatic rain-gauge of Sir Christopher Wren, FRS," *Notes Rec. Royal Soc. Lond.*, vol. 22, no. 1/2, pp. 94–104, 1967.

[3] J. S. Marshall, R. C. Langille, and W. M. Palmer, "Measurement of rainfall by radar," *J. Atmos. Sci.*, vol. 4, pp. 186–192, 1947.

[4] H. Messer, A. Zinevich, and P. Alpert, "Environmental monitoring by wireless communication networks," *Science*, vol. 312, no. 5774, pp. 713–713, 2006.

[5] H. Messer, "Rainfall monitoring using cellular networks," *IEEE Signal Processing Mag.*, vol. 24, no. 3, pp. 144–142, 2007.

[6] H. Leijnse, R. Uijlenhoet, and J. N. M. Stricker, "Rainfall measurement using radio links from cellular communication networks," *Water Resour. Res.*, vol. 43, no. 3, 2007.

[7] D. Giuli, A. Toccafondi, G. B. Gentili, and A. Freni, "Tomographic reconstruction of rainfall fields through microwave attenuation measurements," *J. Appl. Meteorol.*, vol. 30, no. 9, pp. 1323–1340, 1991.

[8] D. Giuli, L. Facheris, and S. Tanelli, "Microwave tomographic inversion technique based on stochastic approach for rainfall fields monitoring," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, no. 5, pp. 2536–2555, 1999.

[9] [Online]. Available: http://prswww.essex.ac.uk/mantissa/

[10] R. Olsen, D. V. Rogers, and D. Hodge, "The aRb relation in the calculation of rain attenuation," *IEEE Trans. Antennas Propagat.*, vol. 26, no. 2, pp. 318–329, 1978.

[11] A. Zinevich, P. Alpert, and H. Messer, "Estimation of rainfall fields using commercial microwave communication networks of variable density," *Adv. Water Resour.*, vol. 31, no. 11, pp. 1470–1480, 2008.

[12] T. L. Frey, "The effects of the atmosphere and weather on the performance of a mm-wave communication link," *Appl. Microw. Wireless*, vol. 11, pp. 76–81, 1999.

[13] A. Zinevich, H. Messer, and P. Alpert, "Prediction of rainfall intensity measurement errors using commercial microwave communication links," *Atmos. Meas. Technol.*, vol. 3, pp. 1385–1402, 2010.

[14] O. Goldshtein, H. Messer, and A. Zinevich, "Rain rate estimation using measurements from commercial telecommunications links," *IEEE Trans. Signal Processing*, vol. 57, no. 4, pp. 1616–1625, 2009.

[15] O. Sendik and H. Messer, "On the reconstructability of images sampled by random line projections," in *Proc. IEEE 27th Convention Institute of Electrical & Electronics Engineers Israel (IEEEI)*, Nov. 2012, pp. 1–5.

[16] H. Leijnse, R. Uijlenhoet, and J. N. M. Stricker, "Hydrometeorological application of a microwave link: 1. Evaporation," *Water Resour. Res.*, vol. 43, no. 4, p. W04416, 2007.

[17] S. J. Birrell and J. W. Hummel, "Multi-sensor ISFET system for soil analysis," *Precision Agri.*, vol. 2, pp. 459–460, 1997.

[18] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proc. 1st ACM Int. Workshop Wireless Sensor Networks Applications*, Sept. 2002, pp. 88–97.

[19] E. M. Trono, M. L. Guico, N. J. C. Libatique, G. L. Tangonan, D. N. B. Baluyot, T. K. R. Cordero, and A. P. F. Parrenas, "Rainfall monitoring using acoustic sensors," in *Proc. TENCON 2012 IEEE Region 10 Conf.*, Nov. 2012, pp. 1–6.

[20] C. G. Panayiotou, D. Fatta, and M. P. Michaelides, "Environmental monitoring using wireless sensor networks," Univ. Cyprus, Dept. of Electrical Engineering, 2005.

[21] P. Corke, T. Wark, R. Jurdak, W. Hu, P. Valencia, and D. Moore, "Environmental wireless sensor networks," *Proc. IEEE*, vol. 98, no. 11, pp. 1903–1917, 2010.

[22] T. Arici and Y. Altunbasak, "Adaptive sensing for environment monitoring using wireless sensor networks," in *Proc. IEEE Wireless Communications Networking Conf. (WCNC 2004)*, vol. 4, Mar. pp. 2347–2352.

[23] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proc. 21st Annu. Joint Conf. IEEE Computer Communications Societies, (INFOCOM 2002)*, vol. 3, pp. 1567–1576.

[24] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. 33rd Annu. Hawaii Int. Conf. System Sciences*, Jan. 2000, p. 10–19.

[25] T. Van Dam and K. Langendoen, "An adaptive energy-efficient MAC protocol for wireless sensor networks," in *Proc. ACM 1st Int. Conf. Embedded Networked Sensor Systems*, Nov. 2003, pp. 171–180.

[26] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Netw.*, vol. 38, no. 4, pp. 393–422, 2002.

[27] A. Zinevich, H. Messer, and P. Alpert, "Frontal rainfall observation by a commercial microwave communication network," *J. Appl. Meteorol. Climatol.*, vol. 48, no. 7, pp. 1317–1334, 2009.

[28] H. Leijnse, R. Uijlenhoet, and J. N. M. Stricker, "Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation," *Adv. Water Resour.*, vol. 31, no. 11, pp. 1481–1493, 2008.

[29] A. Berne and R. Uijlenhoet, "Path-averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability," *Geophys. Res. Lett.*, vol. 34, no. 7, 2007.

[30] A. Zinevich, H. Messer, and P. Alpert, "Prediction of rainfall intensity measurement errors using commercial microwave communication links," *Atmos. Meas. Technol.*, vol. 3, pp. 1385–1402, 2010.

[31] M. M. Z. Kharadly and R. Ross, "Effect of wet antenna attenuation on propagation data statistics," *IEEE Trans. Antennas Propagat.*, vol. 49, no. 8, pp. 1183–1191, 2001.

[32] H. Minda and K. Nakamura, "High temporal resolution path-average rain gauge with 50-GHz band microwave," *J. Atmos. Ocean. Technol.*, vol. 22, no. 2, pp. 165–179, 2005.

[33] M. Schleiss, J. Rieckermann, and A. Berne, "Quantification and modeling of wet-antenna attenuation for commercial microwave links," *IEEE Geosci. Remote Sensing. Lett.*, vol. 10, no. 5, pp. 1195–1199, 2013

[34] A. R. Rahimi, A. R. Holt, G. J. G. Upton, and R. J. Cummings, "Use of dual-frequency microwave links for measuring path-averaged rainfall," *J. Geophys. Res.*, vol. 108, no. D15, p. 4467, 2003.

[35] A. Overeem, H. Leijnse, and R. Uijlenhoet, "Measuring urban rainfall using microwave links from commercial cellular communication networks," *Water Resour. Res.*, vol. 47, no. 12, 2011.

[36] U. Hadar, "High-order HMM with applications to rainfall estimation from cellular network measurements," Ph.D. dissertation, Dept. Electr. Eng., Tel Aviv Univ. 2009.

[37] A. Rayitsfeld, R. Samuels, A. Zinevich, U. Hadar, and P. Alpert, "Comparison of two methodologies for long term rainfall monitoring using a commercial microwave communication system," *Atmos. Res.*, vol. 104, pp. 119–127, 2012.

[38] M. Schleiss and A. Berne, "Identification of dry and rainy periods using telecommunication microwave links," *IEEE Geosci. Remote Sensing. Lett.*, vol. 7, no. 3, pp. 611–615, 2010.

[39] C. Chwala, A. Gmeiner, W. Qiu, S. Hipp, D. Nienaber, U. Siart, and H. Kunstmann, "Precipitation observation using microwave backhaul links in the al-pine and pre-alpine region of Southern Germany," *Hydrol. Earth Syst. Sci.*, vol. 16, no. 8, pp. 2647–2661, 2012.

[40] Z. Wang, M. Schleiss, J. Jaffrain, A. Berne, and J. Rieckermann, "Using Markov switching models to infer dry and rainy periods from telecommunication microwave link signals," *Atmos. Measure. Techn. Discuss.*, vol. 5, pp. 411–445, 2012.

[41] M. Kaufmann and J. Rieckermann. "Identification of dry and rainy periods using telecommunication microwave links," in *Proc. 12th Int. Conf. Urban Drainage*. Porto Alegre, Brazil, 2011, pp. 10–15.

[42] A. R. Holt, J. W. F. Goddard, G. J. G. Upton, M. J. Willis, A. R. Rahimi, P. D. Baxter, and C. G. Collier, "Measurement of rainfall by dual-wavelength microwave attenuation," *Electron. Lett.*, vol. 36, no. 25, pp. 2099–2101, 2000.

[43] N. David, P. Alpert, and H. Messer, "Technical Note: Novel method for water vapour monitoring using wireless communication networks measurements," *Atmos. Chem. Phys.*, vol. 9, no. 7, pp. 2413–2418, 2009.

[44] N. David, P. Alpert, and H. Messer, "The potential of commercial microwave networks to monitor dense fog-feasibility study," *J. Geophys. Res.-Atmos.*, vol. 118, no. 20, pp. 11750–11761, 2013.

[45] K. P. Hunt, J. J. Niemeier, L. K. da Cunha, and A. Kruger, "Using cellular network signal strength to monitor vegetation characteristics," *IEEE Geosci. Remote Sensing. Lett.*, vol. 8, no. 2, pp. 346–349, 2011.

[46] O. Harel and H. Messer, "Extension of the MFLRT to detect an unknown deterministic signal using multiple sensors, applied for precipitation detection," *IEEE Signal Processing Lett.,* vol. 20, no. 10, pp. 945–948, Oct. 2013.

[47] A. Overeem, H. Leijnse, and R. Uijlenhoet, "Country-wide rainfall maps from a commercial cellular telephone network," in *Proc. EGU General Assembly Conf. Abstracts*, vol. 14, Apr. 2012, pp. 5937–5948.

[48] H. Kunstmann, C. Chwala, S. Hipp, W. Qiu, and U. Siart, "Precipitation quantification by cellular-network backhaul-link signal attenuation and a monostatic atmospheric transmission experiment," in *Proc. EGU General Assembly Conf. Abstracts*, vol. 12, May 2010, p. 9444.

[49] A. Overeem, H. Leijnse, and R. Uijlenhoet, "Country-wide rainfall maps from cellular communication networks," *Proc. Natl. Acad. Sci.*, vol. 110, no. 8, pp. 2741–2745, 2013.

[50] L. Barthès and C. Mallet, "On the opportunistic use of geostationary satellite signals to estimate rain rate in the purpose of radar calibration," in *Proc. 7th European Conf. Radar Meteorology Hydrology (ERAD),* 2012.

[51] C. Chwala, H. Kunstmann, S. Hipp, U. Siart, and T. Eibert, "Precipitation observation using commercial microwave communication links," in *Proc. IEEE Int. Geoscience Remote Sensing Symp. (IGARSS),* July 2012, pp. 2922–2925.

[52] J. Ostrometzky, D. Cherkassky, and H. Messer. Accumulated mixed precipitation estimation using measurements from multiple microwave links. *Adv. Meteorol.* [Online]. Available: http://www.hindawi.com/journals/amete/aip/707646/

[53] N. David, P. Alpert, and H. Messer, "Humidity measurements using commercial microwave links," in *Advanced Trends in Wireless Communications*. Rijeka, Croatia: InTech, 2011, pp. 65–78.

[54] D. Cherkassky, Y. Ostrometzky, and H. Messer, "The use of linear feature projection for precipitation classification using measurements from commercial microwave links," in *Latent Variable Analysis and Signal Separation*. Berlin, Germany: Springer, 2012, pp. 511–519.

[55] R. J. Watson and D. D. Hodges, "Estimation of rainfall rate from terrestrial microwave link measurements," in *Proc. IEEE Int. Geoscience Remote Sensing Symp. (IGARSS)*, vol. 3, July 2009, pp. 263–266.

[56] F. Cuccoli, L. Facheris, S. Gori, and L. Baldini, "Retrieving rainfall fields through tomographic processing applied to radio base network signals," *Proc. SPIE*, vol. 8174, pp. 81740C–81740C, Oct. 2011.

[57] M. Grum, S. Kraemer, H. R. Verworn, and A. Redder, "Combined use of point rain gauges, radar, microwave link and level measurements in urban hydrological modelling," *Atmos. Res.*, vol. 77, no. 1, pp. 313–321, 2005.

[58] S. Vogl, P. Laux, W. Qiu, G. Mao, and H. Kunstmann, "Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields," *Hydrol. Earth Syst. Sci.*, vol. 16, no. 7, pp. 2311–2328, 2012.

[59] V. N. Bringi, V. Chandrasekar, N. Balakrishnan, and D. S. Zrnic, "An examination of propagation effects in rainfall on radar measurements at microwave frequencies," *J. Atmos. Oceanic Technol.*, vol. 7, no. 6, pp. 829–840, 1990.

[60] S. I. Ghobrial and S. Sharief, "Microwave attenuation and cross polarization in dust storms," *IEEE Trans. Antennas Propagat.*, vol. 35, no. 4, pp. 418–425, 1987.

[61] J. C. Andrews, "Microwave absorption in compressed carbon monoxide and nitrous oxide," *Aust. J. Phys.*, vol. 25, no. 3, pp. 283–292, 1972.

[62] H. Messer, O. Goldshtein, A. Rayitsfeld, and P. Alpert, "Recent results of rainfall mapping from cellular network measurements," in *Proc. IEEE Int. Conf. Acoustics, Speech Signal Processing (ICASSP),* Mar. 2008, pp. 5157–5160.

[63] N. David, P. Alpert, and H. Messer, "The potential of cellular network infrastructures for sudden rainfall monitoring in dry climate regions," *Atmos. Res.*, vol. 131, pp. 13–21, 2013.

[64] N. David, O. Sendik, H. Messer, and P. Alpert, "Cellular network infrastructure: The future of fog monitoring?" *Bulletin of the American Meteorological Society*, doi 10.1175/BAMS-D-13-00292.1, 2014.

[SP]

[
Jarmo Lundén,
Visa Koivunen,
and H. Vincent Poor
]

Spectrum Exploration and Exploitation for Cognitive Radio

[Recent advances]

The lack of availability of radio spectrum for wireless communication purposes is becoming a serious problem as more wireless systems and services are being developed and operate in crowded spectral bands. The scarcity of useful radio spectrum is mainly due to the static allocation and rigid regulation of the spectrum use rather than the spectrum being actually fully in use. Flexible spectrum use and cognitive radio technologies provide an approach to alleviating this problem by allowing for secondary spectrum use while the spectrum is underutilized by its primary licensed users. Idle spectrum is a time–frequency–location varying resource. It is a resource that also depends on the relative locations of the primary and secondary receivers and transmitters as well as the instantaneous propagation conditions. By acquiring awareness about the current radio environment and the other spectrum users, cognitive radios can more efficiently exploit idle spectrum and manage interference. Doing so requires a means to explore the spectrum to identify high-quality and persistent local spectral resources and access and share them among a number of users while strictly controlling the interference caused to others, in particular, licensed primary users (PUs). Situational awareness about the state of the spectrum allows for optimal exploitation of underutilized spectrum. For example, idle subbands may be allocated, and waveform parameters may be chosen to maximize the sum-rate for the cognitive users while making sure no harmful interference is caused to the other users of the spectrum.

The purpose of this article is to present recent advances in spectrum exploration and exploitation. The goal is to jointly optimize the identification and access of underutilized spectrum in multiband and multiuser environments where the state of the spectrum may vary rapidly. Cognitive users may acquire the necessary information about the state of the spectrum by sensing it and sharing the results among the users. Reviews of spectrum sensing can be found in [4] and [5]. Spectrum sensing is a key enabler of spectrum exploration through which situational awareness about the spectrum is developed and the behavior of the PU traffic is learned as a

function of time, location, and frequency. Spectrum exploration and exploitation over multiple frequency bands may be considered an optimization or machine-learning problem. A reward or utility function is maximized to guide the exploration and exploitation such that the idle spectrum is used as efficiently as possible while managing the interference. In cooperative multiband and multiuser scenarios, spectrum sensing and access policies allocate and divide the sensing, learning, and interference management tasks among the network nodes and different subbands of the spectrum so that idle spectrum may be optimally identified and exploited. In noncooperative competitive environments, the network nodes compete for the available spectrum to maximize their own utilities. An acquired awareness of the state of the spectrum may be used for scheduling, waveform selection, and radio resource management purposes as well.

ADVANCED SPECTRUM SENSING TECHNIQUES

Spectrum sensing using a single sensor to identify idle spectrum is a widely addressed topic in signal processing and wireless communications research; see [4] and [5]. Commonly used methods include energy detection and methods relying on statistical or structural properties of communications waveforms such as cyclostationarity (see "Cyclostationarity-Based Spectrum Sensing"), low-rank signal structure, or statistics of eigenvalues of a correlation matrix.

Moreover, the autocorrelation structure induced by the modulation scheme, such as the cyclic prefix in orthogonal frequency division multiplexing (OFDM), and known pilot waveforms have been used. Wireless standards typically specify all of the employed waveforms in detail. Efficient methods used for signal detection or classification purposes may be derived based on such specifications. Spectrum sensing may also collect valuable information to create awareness of the state of the spectrum. Spectrum awareness may be used for interference modeling and management, power control, scheduling, resource allocation, and routing purposes.

In this article, we focus on finding and accessing idle spectrum in multiband and multiuser scenarios. Multiple sensors or cognitive radios in different locations sense multiple subbands to identify idle bands that may be used for data transmission. Not all of the radios should sense the same band at the same time. Instead, the sensing tasks need to be allocated among the users and bands in an efficient manner to speed up the sensing and provide the desired diversity gains. After an idle band is identified, the cognitive radios have to decide who gets access to the unoccupied spectrum and whether it should be accessed at all if low-quality channels are observed or constraints imposed on harmful interference caused to the other users may be violated. We first consider the multiuser spectrum sensing problem.

CYCLOSTATIONARITY-BASED SPECTRUM SENSING

Cyclostationarity-based distributed spectrum sensing allows for distinguishing among secondary user (SU) and PU waveforms exhibiting cyclostationarity at different cycle frequencies, relaxes the assumptions on noise statistics, and exhibits a reliable performance in difficult propagation environments. Cooperative cyclostationarity-based detection algorithms based on generalized log-likelihood ratio (GLLR) have been proposed in [33]. The proposed local multicycle detectors are based on testing whether the cyclic autocorrelation of the received signal is nonzero at the cycle frequencies of interest. The cyclic autocorrelation of a received complex-valued signal $y(k)$ at cycle frequency α and delay τ may be estimated as

$$\hat{R}_y^\alpha(\tau) = \frac{1}{N_s} \sum_{k=1}^{N_s} y(k) y^*(k+\tau) e^{-j2\pi\alpha k}, \qquad (S1)$$

where N_s is the number of observations.

The local GLLR test is given by [33]

$$\mathcal{T} = N_s \hat{\mathbf{r}}_y \hat{\Sigma}_y^{-1} \hat{\mathbf{r}}_y^T \underset{H_0}{\overset{H_1}{\gtrless}} \xi, \qquad (S2)$$

where ξ is the test threshold, $\hat{\mathbf{r}}_y$ is a $1 \times 2M$ vector consisting of stacked real and imaginary parts of estimated cyclic autocorrelations $\hat{R}_y^\alpha(\tau)$ at the cycle frequencies of interest $\alpha \in \mathcal{A}$ for a set of time lags $\{\tau_{i,j}\}_{i=1}^{N_{\tau,j}}$ for each $\alpha_j \in \mathcal{A}$, and $\hat{\Sigma}_y$ is an estimate of the asymptotic covariance matrix of $\hat{\mathbf{r}}_y$ that can be obtained, e.g., using frequency-smoothed cyclic periodograms. Thus, $M = \sum_{j=1}^{|\mathcal{A}|} N_{\tau,j}$.

In a multiuser scenario with independent users, the GLLR at a fusion center (FC) is the sum of the local GLLRs $\mathcal{T}_N = \sum_{k=1}^N \mathcal{T}_i$,

where $\mathcal{T}_i$ is the GLLR of SU i. This sensing approach may be made more energy efficient by using censoring. The SUs communicate only those GLLRs exceeding the censoring threshold ψ_i to the FC as described in (4) and (5). Figure S1 illustrates that censoring combined with the multicycle detector in (S2) results in only a small performance loss even with very strict communication rate constraints.
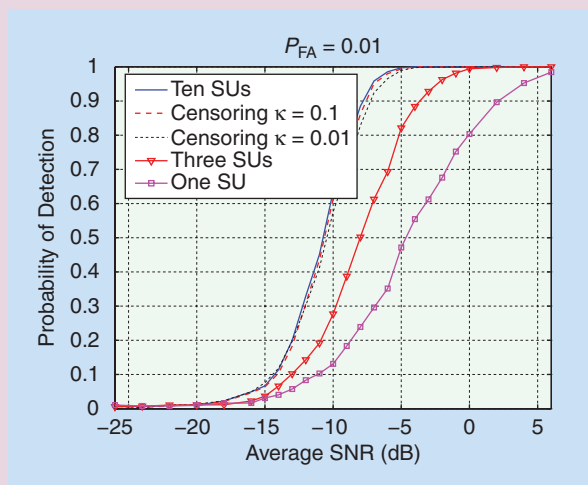
[FIGS1] The probability of detection versus the average SNR (dB). This is done to detect an OFDM signal in a Rayleigh fading channel with log-normal shadowing using cyclostationarity-based local detectors [33] and different communication rate constraints for SUs communicating sensing results to the FC.

OVERVIEW OF DISTRIBUTED SENSING

Distributed sensing involving multiple geographically displaced sensors and an FC has been studied in the sensor network and radar research communities [47], [60]. Recently, it has found applications to cooperative spectrum sensing in cognitive radio systems; see, e.g., [9], [21], [26], [33], and [68]. Distributed spectrum sensing is particularly suitable for creating awareness of the state of the radio environment since idle spectrum is a local resource that varies as a function of time, location, and frequency. Understanding the state of the spectrum is crucial for interference management, scheduling, and resource allocation.

Cooperation in multiuser spectrum sensing has several advantages. It provides diversity gains in the face of demanding propagation environments such as fading, shadowing, and the hidden node problem. Various propagation effects make it harder to detect the signals, as the signal attenuation caused by them may be of the order of tens of decibels. The gains obtained by cooperative sensing follow from spatial diversity in the same way as in multiantenna (multiple-input, multiple-output) wireless communications. If the sensor displacement is sufficiently large, the sensing channels can be assumed to be independent from sensor to sensor. Distributed sensing leads to improvement in the detector performance in terms of fewer false alarms (type I errors) and missed detections (type II errors) and a shorter detection time at the specified performance level. Multiple sensors also provide inherent redundancy, which leads to higher reliability and robustness. However, by increasing the number of cooperating sensing nodes, the amount of overhead traffic will increase and the amount of system resources needed in fusing the sensing results increases as well. Thus, one may be able to determine an optimum number of cooperating sensors, beyond which the sensing performance improvement does not compensate for the increased use of resources [11], [36]. Optimizing the use of sensing resources has also been considered, e.g., in [25] and [44].

Mobile terminals intended for agile spectrum use may have a spectrum sensing capability. Then, multiple devices can form a distributed sensing system. Cooperation facilitates the use of a simpler and more energy-efficient sensor in each node. Hence, an extended battery life may be achieved. On the other hand, exchanging control information and sharing the sensing results with the FC or the other nodes increases the energy consumption in the network. Furthermore, policies used for sensing and accessing the spectrum also play a crucial role in determining the overall power consumption.

There are many ways to characterize the gains obtained through spatial diversity. A suitable quantitative measure for cooperative detection in cognitive radio networks is the probability of detection P_D as a function of the logarithm of the signal-to-noise ratio (SNR) ρ for different numbers of cooperating sensors. We may define the spatial diversity order (D) as [41]

$$D(N) = \max \frac{\partial}{\partial \rho_{dB}} P_D(\rho_{dB}, N), \qquad (1)$$

where $\rho_{dB} \triangleq 10\log_{10}\rho$, $\partial/(\partial\rho_{dB})$ denotes the partial derivative with respect to (w.r.t.) ρ_{dB}, and N is the number of sensors. Figure 1 depicts the spatial diversity as a function of the number
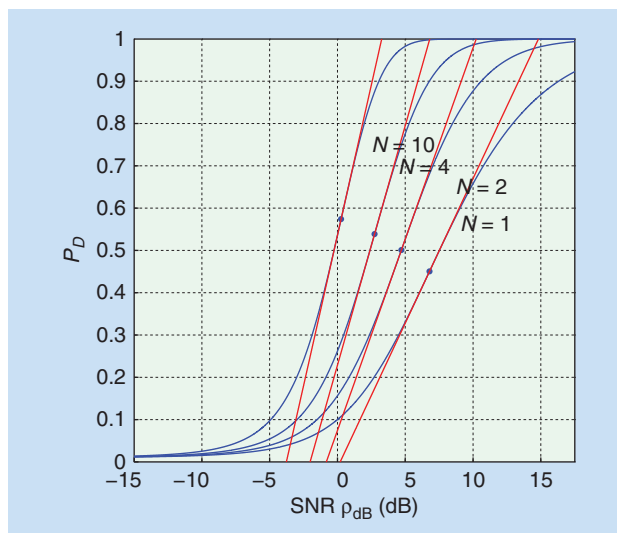
[FIG1] The spatial diversity order may be defined as the maximum slope of the probability of detection curve as a function of the logarithmic SNR ρ_{dB} [41]. In this case, there are N i.i.d. Rayleigh channels, and the employed detection scheme is energy detection with $P_{FA} = 0.01$.

of cooperating sensors for cooperative energy detection (radiometry) using equal-gain combining (EGC) in independent and identically distributed (i.i.d.) Rayleigh fading channels. It can be clearly seen from the figure how the slope of the probability of detection curve grows as the number of cooperating nodes sensing the same band increases. Adding more sensors to observe a subband yields diminishing returns, however. The achieved gain also depends on the fusion rule used in combining the local sensing information as well as the local propagation conditions.

CENTRALIZED OR DECENTRALIZED PROCESSING

Distributed sensing systems can be either centralized or decentralized. In a centralized system, all of the local sensors transmit all of the locally observed data to a central node that performs the data fusion needed for detection or parameter estimation tasks. The sensors in a decentralized system are capable of processing the data locally before transmitting the results to an FC. Local hard (binary) decisions or some other compact statistic, e.g., a sufficient statistic, characterizing the state of the spectrum may be sent to the FC. Consequently, the FC of a decentralized system has only part of the information received from the sensors, whereas, in a centralized system, the FC has all of the information. The channel capacity requirements of a centralized scheme in which all sensing results are transmitted to the central processor may be prohibitive in practice. Therefore, we focus on decentralized distributed spectrum sensing for cognitive radios. In such systems, the nodes typically collaborate in identifying and exploiting the underutilized spectrum.

In distributed sensing systems, the most common topologies with an FC are serial, parallel, and tree topologies [60]. The serial topology is not robust since a single link failure or missing node will cause severe performance degradation. In the parallel topology, the sensors do not typically communicate with each other, and no feedback is provided from the FC to the sensors. The sensors use mapping

rules $u_i = \nu_i(y_i)$, $i = 1, \ldots, N$, where $y_i = [y_{i,1}, \ldots, y_{i,N_S}]^T$ is an N_S-dimensional vector of the observations made by sensor i and N is the number of sensors, and then pass the local mappings, u_i, to the FC. The mapping compresses the data, for example, by sending a local binary decision to the FC. Sensing nodes that are close to each other may also form a cluster that sends summary information about the state of the spectrum to the FC. In an ad hoc configuration, there is no dedicated FC and the sensing information is distributed to all of the nodes. Hence, each node will have the same information needed to decide if the spectrum is idle. Distributing the information to all nodes potentially over multiple hops may cause unacceptable delays in decision making. Similar robustness to node failures is obtained with consensus algorithms [51], [64] in which the sensors exchange information locally over multiple rounds to reach a consensus on some global function of the data, such as a global decision on spectrum occupancy. A drawback to this approach is its iterative nature, which is both time and energy consuming.

The distributed spectrum sensing problem is typically modeled as a binary hypothesis test. The selected decision-making strategy, such as Neyman–Pearson, minimax, or Bayes, is also related to controlling interference. If missed detections occur frequently, there will be collisions with the primary signal and, as a result, retransmissions by both the PUs and the SUs are needed. Many false alarms mean that the opportunities to use idle spectrum are overlooked. Therefore, the improvements in spectrum efficiency obtained by the secondary spectrum use are reduced. As an example, the Neyman–Pearson scheme for distributed detection can be stated as follows: For a predefined global probability of false alarm level P_{FA}, find (optimum) local and global decision rules $\Gamma = \{\nu_0(u), \nu_1(y_1), \ldots, \nu_N(y_N)\}$, where $\nu_0(\cdot)$ is the global decision rule and $\nu_i(\cdot), i = 1, \ldots, N$ are the local rules that minimize the global probability of missed detection P_{MD}.

A key assumption that facilitates finding a tractable solution to the distributed detection problem is the conditional independence of observed sensor data conditioned on the hypothesis. If this assumption holds, the local mapping rules as well as the global decision rule at the FC become likelihood-ratio-based threshold rules; see [60]. The null hypothesis H_0 at the FC is that the spectrum is idle (noise only), and the alternative hypothesis H_1 is that a primary signal is present. The decision between these two is made by comparing the likelihood ratio $\Lambda(u)$ to a threshold value ψ_0. From conditional independence, it follows that $\Lambda(u)$ may be given in a factored form

$$\Lambda(u) = \prod_{i=1}^{N} \frac{p(u_i \mid H_1)}{p(u_i \mid H_0)}, \qquad (2)$$

where $p(\cdot)$ is used to denote probability densities or mass functions of its argument(s). However, if the conditional independence assumption is not valid, the optimal tests are no longer simple threshold rules based on the likelihood ratios at the individual sensors. Conditionally dependent sensor data could arise, for example, because of correlated shadowing if the local sensors are in close proximity to each other. A comprehensive treatment of shadowing and other propagation issues in cognitive radio networks can be found in [5, Ch. 3].

If the local decision variables are binary, the global decision rule is a Boolean rule and the log-likelihood ratio (LLR) at the FC can be written as a weighted sum of local sensor decisions

$$\log \Lambda(u) = \sum_{i=1}^{N} \left[u_i \log \frac{1 - P_{MD,i}}{P_{FA,i}} + (1 - u_i) \log \frac{P_{MD,i}}{1 - P_{FA,i}} \right], \quad (3)$$

where u_i is the binary local decision of the ith sensor (i.e., 0 or 1), $P_{FA,i}$ is the probability of false alarm at the ith sensor, and $P_{MD,i}$ is the probability of missed detection at the ith sensor. The weights depend on the performance of individual sensors that may not be known in practice. Computationally simpler decision rules are obtained by using well-known Boolean K-out-of-N rules such as OR, AND, and MAJORITY. The detector is commonly designed so that the levels of P_{FA} and P_{MD} at the FC are controlled; see [60].

The local sensors may also compress the local observations by calculating and transmitting the local likelihood ratios. Under the conditional independence assumption, the optimal global decision rule is a likelihood ratio test and the global likelihood ratio is the product of the local likelihood ratios. In practice, the likelihood ratio may include unknown nuisance parameters. Hence, a generalized likelihood ratio (GLR) test obtained by replacing the unknown parameter values with their estimates is often used. The resulting test statistic at the FC is simply the product of the local GLRs. Also, belief propagation techniques have been employed in combining. Two commonly employed linear combining schemes for cooperative detection are EGC and maximal ratio combining (MRC) [35]. The MRC-based fusion scheme is attractive in the low SNR regime, while the EGC-based fusion scheme is a good choice for combination under limited knowledge of channel state information.

In wireless networks, user terminals are typically battery operated. Thus, low power consumption is a desirable design goal. In spectrum sensing, the power consumption depends obviously on the employed sensing algorithms and their implementation including the radio frequency-intermediate frequency front end, the duty cycle used in sensing, the employed sensing policy, and the power used in reporting the sensing results. In addition, deciding if and how the idle spectrum is accessed plays a key role in prolonging battery life. For example, if the SU is experiencing poor channel quality even if the channel is decidedly idle, it may make sense not to access the channel for transmitting with a low rate while using high transmit power.

Censoring reduces power consumption by sending only sufficiently informative decision statistics to the FC [47]. An SU sends its test statistic to the FC only when its test statistic, denoted here by $\log \Lambda_i$, is above a censoring threshold defined by the communication rate constraint specified by the designers

$$P(\log \Lambda_i > \psi_i \mid H_0) \leq \kappa_i, i = 1, \ldots, N, \qquad (4)$$

where $\kappa_i \leq 1$ is the communication rate of user i and ψ_i is the upper limit of the censoring (no-send) region of the user i. The ψ_i is chosen such that the probability of user i transmitting the test statistic to the FC under H_0 is κ_i. In the cooperative sensing context, we may use the following censoring test statistic (L-out-of-N users transmit) [33]:

$$\mathcal{D}_N = \sum_{i=1}^{L} \log \Lambda_i + \sum_{i=L+1}^{N} d_i, \tag{5}$$

where it is assumed that the first L users with sufficiently informative decision statistics transmit and $d_i = E\left[\log \Lambda_i \mid \log \Lambda_i \leq \psi_i, H_0\right]$ is the conditional mean of the local LLR of the ith SU in the no-send region under the null hypothesis. The no-send region is not ignored but captured by a single quantity, i.e., the conditional mean of the log-likelihood in the no-send region, which is optimal in the minimum mean-square error sense. In practice, there is no need to transmit d_i as it can be calculated at the FC since the limits ψ_i are determined by the communication rate constraints κ_i.

Using binary (hard) local decisions in a distributed decentralized detection system reduces the communication cost at the expense of loss of information. Using soft decisions, such as LLRs or their quantized versions, typically leads to an improvement in performance. It is commonly argued that using soft decisions significantly increases the amount of data to be transmitted and, hence, the power consumption. This is not necessarily true since there may be significant overhead related to frame structure and different layers of the communication protocol stack used in transmitting the sensing results to the FC. That overhead will be present even if only binary decisions are transmitted. As a result, the difference in the transmission and bandwidth requirements between hard and soft decision statistics may be small. Individual sensors may also convey information about interference levels, channel quality, and occupancy to build awareness of the state of the spectrum.

The analytical studies in [10] indicate that by using ≥ 4 bits for the quantization of the likelihood ratios, the performance loss remains negligible. Error-free reporting channels and a local autocorrelation-based detector exploiting the cyclic prefix of the OFDM modulation are assumed in [10]. The impact of quantization on distributed detection has also been analyzed in [60]. Reporting channels may introduce errors. Powerful channel coding and higher transmit power levels are obvious solutions to reduce these. However, this may not be practical if low power consumption is needed. The impact of reporting channel errors in cooperative sensing has been studied in [10] and [26]. A phenomenon known as the *bit-error probability* (*BEP*) *wall* was reported in [10] for both hard- and soft-decision-based cooperative sensing. If the BEP of the reporting channel exceeds the BEP wall value, then, irrespective of the received signal quality on the listening channel or the sensing time at the SU, constraints on the detector performance cannot be met at the FC. Soft-decision-based systems are more robust in the face of reporting channel errors.

SEQUENTIAL SENSING METHODS

Sequential and quickest detection techniques have applications in a number of fields, such as radar, fault detection, finance, and clinical trials, among others [23], [46]. In cognitive radio systems, sequential detection (SD) techniques are important for detecting changes in the state of the spectrum, i.e., rapidly identifying new spectral opportunities as well as vacating a specific frequency band quickly when the PU becomes active. Sensing time is an important parameter in finding idle spectrum. Reducing the time spent for sensing allows the time used for transmission to be increased. Furthermore, fewer energy resources are spent for sensing. SD aims at minimizing the detection time for a desired performance level. In a nonsequential test, the sample size is fixed, whereas, in the case of a sequential test, it varies depending on the data. In cognitive radio systems, SD may be used for both single-user and collaborative distributed detection tasks. Moreover, in the collaborative case, SD algorithms may be used either only at the local sensors, the FC, or both.

We will briefly consider two sequential analysis approaches. Classical SD aims at distinguishing between two hypotheses from a sequence of i.i.d. random observations. The objective is to make a decision as quickly as possible given specified error levels. An alternative formulation is the quickest detection problem in which the objective is to detect a change in the distribution of the data, i.e., find the change point, with minimal detection delay.

Let $y_1, y_2, \ldots$, be a sequence of i.i.d. random observations with a common distribution F_0 or F_1. The binary test of hypothesis H_0 against H_1 may be formulated as

$$\begin{aligned} H_0 &: \ y_n \sim F_0, n = 1, 2, \ldots, \\ H_1 &: \ y_n \sim F_1, n = 1, 2, \ldots. \end{aligned} \tag{6}$$

Let p_0 and p_1 denote the probability density functions associated with F_0 and F_1, respectively. SD aims at choosing between these two hypotheses in a way that minimizes the number of observations given constraints on the type I or II errors. The sequential probability ratio test (SPRT) of Wald requires the minimal average number of observations under both hypotheses among all tests with equal (or smaller) error probabilities [61]. The stopping time of the SPRT is given by

$$N_S = \inf\{n \geq 1 \mid \Lambda_n \leq B \text{ or } \Lambda_n \geq A\}, \tag{7}$$

where $\Lambda_n = \prod_{k=1}^{n} p_1(y_k)/p_0(y_k)$ is the likelihood ratio and A and B are upper and lower stopping boundaries, respectively. Thus, after each sample, the SPRT accepts H_1 if $\Lambda_n \geq A$, or accepts H_0 if $\Lambda_n \leq B$. If $B < \Lambda_n < A$, it takes an additional observation. The stopping boundaries A and B may be chosen based on the target levels of the probability of false alarm and the probability of missed detection, respectively. The idea of the SPRT is illustrated in Figure 2.

The hypotheses have to be simple and the distributions completely specified under both hypotheses for the SPRT to be optimal. In cognitive radio applications, however, there are often unknown nuisance parameters related to transmit powers, propagation conditions, adaptive modulation and coding schemes employed, and different PU modes. Thus, there have been numerous efforts to design SD tests for the case of composite hypotheses, such as sequential GLR tests [23] and minimax tests [6]. Moreover, it is often necessary to ensure that a decision is reached within a certain time frame. Hence, a truncated test [57] may have to be employed. A truncated test uses a final decision rule to decide between the two hypotheses if a predefined upper limit on the number of observations is reached before a decision has been made by the SPRT.
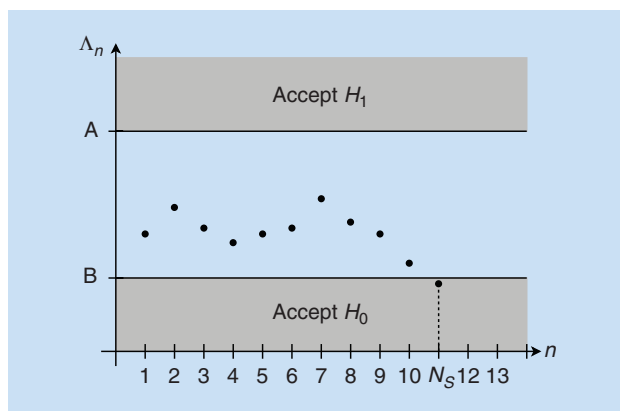
[FIG2] An illustration of the SPRT.

A decentralized SPRT (D-SPRT) scheme in which both the local sensors and the FC employ SPRTs is proposed in [18]. The local sensors employ repeated SPRTs. After making a local decision (i.e., the SPRT stops), the local detector transmits the binary decision to the FC and starts a new local SPRT on its subsequent observations. The FC employs an SPRT on the received local decisions. Figure 3 illustrates the benefit of sequential testing in a decentralized system of [9] in comparison to fixed-sample-size (FSS) testing by plotting the number of SU statistics used at the FC versus SNR. An SPRT at the FC is used, and each local sensor employs an autocorrelation-based detector for OFDM signals [9]. The SUs transmit their local LLRs to the FC. The probability of false alarm is set to $\alpha = 0.05$, and the probability of missed detection is set to $\beta = 1 - P_D = 0.05$. Transmission is assumed to take place over an additive white Gaussian noise (AWGN) channel.

QUICKEST DETECTION

Let $y_1, y_2, \ldots,$ be a sequence of independent random observations with an unknown change point m, such that $y_1, \ldots, y_{m-1}$ have a
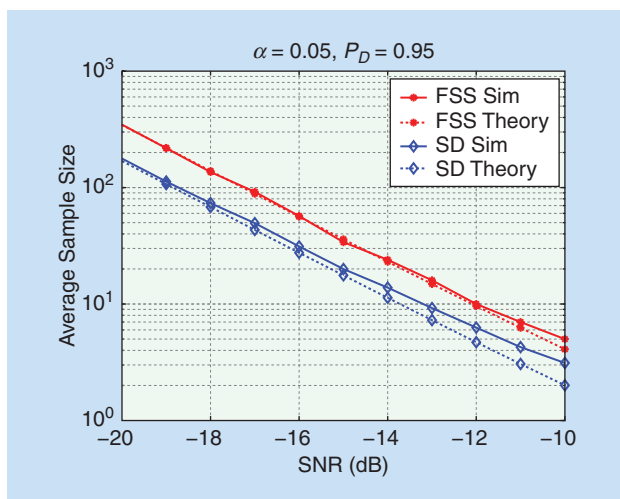
common distribution F_0 and $y_m, y_{m+1}, \ldots$ have another common distribution F_1. Let p_0 and p_1 denote probability density functions of F_0 and F_1, respectively. Both Bayesian and non-Bayesian quickest detection approaches have received considerable attention in the research community; see [46]. Page's cumulative sum (CUSUM) test [43] is the most commonly used non-Bayesian quickest detection method. The stopping time of the CUSUM test for detecting a change is given by

$$N_S = \inf\{n \geq 1 \mid U_n \geq \psi\}, \qquad (8)$$

where $U_n = \max_{1 \leq k \leq n} \prod_{i=k}^{n} p_1(y_i)/p_0(y_i)$ and $\psi \geq 0$ is the test threshold. The CUSUM test statistic U_n may be updated recursively via $U_n = \max\{1, U_{n-1}\}p_1(y_n)/p_0(y_n), n \geq 1, U_0 = 1.$ The threshold value may be chosen appropriately depending on the detection strategy, for example, to obtain optimality in the minimax sense [37]. Page's CUSUM scheme considers the change point m to be an unknown constant. An alternative Bayesian formulation is obtained by treating the change point as a random parameter with a prior distribution [53].

In cognitive radio applications, it is often unrealistic to assume exact knowledge of the distributions under both hypotheses. The optimum quickest detection algorithms are particularly sensitive to uncertainty in distribution parameters. Many of the approaches dealing with unknown distribution parameters are based either on nonparametric approaches [45] or GLR-based algorithms [23] or are derived assuming least favorable pre- and postchange distributions [58].

A variety of single-user quickest detection methods for cognitive radio systems have been developed; see, e.g., [22]. Multiband scenarios are considered in [67]. In [27], the goal is to detect the new activity of PUs and to choose the best frequency band to sense. A tradeoff between the minimization of false alarms and detection delays is found. Dynamic programming is employed to obtain a control policy for selecting which frequency band to sense and when to declare that a PU has become active.

Collaborative quickest detection schemes for cognitive radio systems have been introduced, e.g., in [15] and [59]. Each local sensor uses the CUSUM algorithm for change detection [15]. The sensors communicate with the FC only after a change in the distribution is detected, and the FC declares a change after receiving at least one message from the sensors. This approach has been shown to be asymptotically optimal as the mean time between false alarms tends to infinity.

OPTIMIZATION OF SPECTRUM EXPLORATION AND EXPLOITATION

We will now provide an overview of optimized spectrum exploration and exploitation. This problem can be formulated as a sensing and access policy design problem. We will begin by first formulating the single- and multiuser spectrum sensing and access problems. Then, we will describe various approaches to solving these problems. The described techniques are dynamic programming, bandit problems, reinforcement learning, and game-theoretic approaches.

[FIG3] A comparison of the theoretical average number of SU decision statistics for SD using the SPRT at the FC and FSS detection needed for equal performance [9]. The sequential method requires significantly fewer decision statistics than FSS detection for the same error probabilities.

SINGLE-USER SENSING AND ACCESS PROBLEM FORMULATION: MARKOV DECISION PROCESSES

In this article, we consider multiband spectrum sensing and access problems in which the spectrum of interest may be extremely wide and noncontiguous. Hence, we consider that each SU can sense and access only part of the spectrum at each time. In such single-user multiband spectrum sensing and access problems, the decision maker, e.g., an SU or FC, has to decide which frequency bands to sense or access at each time. These are stochastic sequential decision problems that can be modeled as Markov decision processes (MDPs). A finite MDP consists of

- a sequence of discrete time steps $n = 0, 1, 2, \ldots$
- a decision maker i
- a finite set of possible states of the environment $s \in S$
- a finite set of possible actions in each state $a \in A$
- a state transition function $\varphi : S \times A \times S \to [0, 1]$ that defines the transition probability $P(s_{n+1} \mid s_n, a_n)$
- a reward/payoff function $r : S \times A \times S \to \mathbb{R}$ that gives reward/payoff r_{n+1} for taking action a_n in state s_n resulting in new state s_{n+1}.

As an example, a multiband and multiuser spectrum sensing problem could be modeled as an MDP as follows. The decision maker is the FC. The FC chooses which frequency band each SU senses at each time. Thus, actions correspond to sensing a particular set of frequency bands by the SUs. The states of the environment would be formed by the PU occupancy of the frequency bands. Thus, if each frequency band can be either vacant or occupied, there are in total 2^{N_B} different states, where N_B is the number of different frequency bands. Finally, the FC would get a reward equal to one for each frequency band sensed vacant and zero for the other frequency bands. Note that this example is just one way of formulating this problem. Many different variations of this basic formulation could be obtained by defining the variables differently. For example, assuming that there is feedback from the spectrum access, the rewards could depend on the obtained throughput or the action space could be expanded by allowing an idle action to be chosen to conserve the SUs' batteries in mobile applications or by allowing the sensing time to be chosen.

The goal of the decision maker is to find an optimal policy π^* that determines in each state $s \in S$ the optimal action $a \in A$ so that a particular objective function is maximized; see Figure 4. A policy π is, thus, a stochastic or deterministic function that determines how the decision maker selects its action a_n in each state s_n. Taking the action a_n in state s_n results in a new state s_{n+1} and a scalar reward r_{n+1} for the agent. A suitable objective function for spectrum sensing and access problems is the expected sum of discounted rewards over an infinite horizon given an initial state $s_n = s$

$$J_n = E\left[\sum_{k=0}^{\infty} \gamma^k r_{n+k+1} \middle| s_n = s \right], \qquad (9)$$

where r_n is the reward at time n and γ, $0 \leq \gamma < 1$, is the discount rate. This objective function gives decreasing weight to future rewards. Other possible objective functions include (discounted) expected reward over a finite time horizon, average
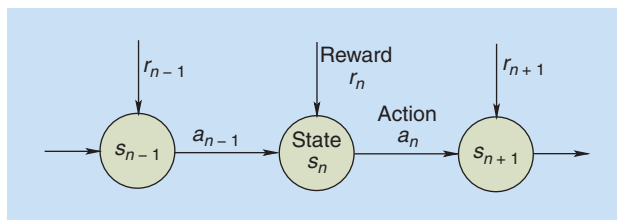
[FIG4] A Markov decision process. In each state s_n, the decision maker chooses an action a_n that results in reward r_{n+1} and new state s_{n+1}. The goal of the decision maker is to maximize a given function of the rewards, such as the expected sum of discounted rewards over a finite or infinite horizon.

expected reward over an infinite time horizon, and regret [24], which measures the expected loss compared to an optimal policy. Returning to our previous example, MDP formulation of a multiband and multiuser spectrum sensing problem, the quantity (9) would be the expected sum of the discounted number of frequency bands found vacant or the expected sum of obtained throughputs over the frequency bands if feedback is available and would place more importance on finding vacant frequency bands or maximizing the obtained throughput at the present time than in the more distant future.

In general, the state transition probabilities are unknown in spectrum sensing and access problems. Moreover, because of limited sensing resources, the spectrum state may not be fully observable either. The partially observable MDP (POMDP) [20] is a generalization of the MDP in which the decision maker cannot directly observe the state. Instead, the decision maker receives observations that depend on the state through some stochastic function. In spectrum sensing problems, instead of observing the state directly, the SUs observe the PU state through sensing results subject to errors.

MULTIUSER SENSING AND ACCESS PROBLEM FORMULATION: GAME THEORY

Game theory can be used to model the interaction of multiple users in spectrum sensing and access problems. Game theory provides models for both noncooperative and cooperative users. In fact, game theory can be divided into noncooperative and cooperative game theory. This categorization is, however, somewhat misleading since cooperation is allowed in both noncooperative and cooperative games. The main difference between noncooperative and cooperative games is that, in noncooperative games, the players act independently and cooperation cannot be enforced, while, in cooperative games, the players act as groups and cooperation within the group can be enforced. Hence, noncooperative games are suited for scenarios in which the cognitive radio users compete with each other and make decisions independently, while cooperative games fit scenarios with cooperative users that aim to jointly optimize the spectrum utilization. Therefore, both of these games have uses in modeling particular multiuser spectrum sensing and access problems.

NONCOOPERATIVE GAMES

In spectrum sensing and access problems, noncooperative games are ideally suited to scenarios in which the SUs operate

independently and are interested only on maximizing their own throughputs. Thus, the SUs are competing for the available spectrum and would generally cooperate with other users only if it improves their own utilities and throughputs.

The most basic model for the interaction of multiple decision makers is a strategic game that consists of

- a set of players $\mathcal{N} = \{1, \ldots, N\}$
- a set of strategies (actions) for each player A_i, $i \in \mathcal{N}$; a strategy $a_i \in A_i$ is a complete plan of action for each situation in the game; the combined strategy space is the set of strategy profiles $A = A_1 \times \cdots \times A_N$
- reward/payoff functions $r_i : A \to \mathbb{R}$, $i \in \mathcal{N}$, that give the players rewards/payoffs r_i for the joint strategies $a_1, \ldots, a_N$.

For example, a multiband and multiuser spectrum access problem could be formulated as a strategic game as follows. The SUs would be the players in the game. The set of strategies could involve the choice of frequency band to access or, for example, the employed transmit power on each frequency band. The reward could then be each SU's individual sum throughput over the different frequency bands.

The objective of each player is to find a strategy that maximizes its reward r_i. The Nash equilibrium is a central concept in noncooperative game theory for establishing the outcome of a game. The Nash equilibrium defines each player's best response strategy given the other players' strategies, i.e., a Nash equilibrium is defined as a strategy profile $a^* \in A$ such that

$$ r_i(a_i^*, a_{-i}^*) \geq r_i(a_i, a_{-i}^*), \forall i, a_i \in A_i, \tag{10} $$

where a_{-i} denotes the strategies of all players except player i. The Nash equilibrium states that a player cannot improve its reward by unilaterally changing its strategy if the other players follow the Nash equilibrium strategies.

A player's strategy may be either a pure or mixed strategy. A pure strategy is a deterministic strategy determining the action in any possible situation. A mixed strategy assigns a probability to each pure strategy, and, thus, the player chooses randomly which pure strategy to play. Every game with finitely many actions has at least one mixed-strategy Nash equilibrium [38]. Since there may be multiple equilibria, finding the best one is important. However, because of the competitive nature of the game, defining the optimality criterion for an equilibrium is not straightforward. One such criterion is Pareto optimality. An equilibrium is Pareto optimal if there exists no other strategy profile that would increase at least one player's reward without decreasing any other player's reward. For more information on Pareto optimality and other criteria and techniques for choosing an equilibrium and improving inefficient equilibria, see [31].

NONCOOPERATIVE GAMES FOR SPECTRUM SENSING AND ACCESS

In [52], a class of nonconvex, noncooperative games is proposed for a multiuser and multiband spectrum sensing and access problem. In the proposed formulation, the SUs compete to maximize their own throughputs by jointly choosing their sensing durations, detection thresholds, and power allocations on the frequency bands. To facilitate decentralized optimization with global PU interference constraints, a pricing mechanism is introduced that penalizes the SUs for their contributions to the total interference. Sufficient conditions for the existence and uniqueness of a Nash equilibrium are derived, and distributed algorithms are proposed for solving the games.

COOPERATIVE GAMES

In general, if the players are interested in maximizing the mutual payoff instead of maximizing only their individual payoffs, cooperation among the players has the potential of improving the overall cumulative payoff compared to that obtained in noncooperative games. In cognitive radio systems, the SUs may cooperate in various ways: the users may share their local sensing results to make finding idle spectrum more efficient, or they may coordinate their sensing and access choices to improve the overall throughput. In addition, the users may help each other, e.g., in routing, packet forwarding, and interference management. Coalitional games constitute one of the most common and important forms of cooperative games.

Coalitional games are cooperative games in which the players form coalitions to improve their rewards. A coalition is a group of players that may enforce cooperation within the coalition. A coalitional game can be seen as a game between coalitions instead of between the individual players. A coalitional game consists of a finite set of players $\mathcal{N} = \{1, \ldots, N\}$ and a coalition value function v that quantifies the value of each coalition $\mathcal{S} \subseteq \mathcal{N}$.

The value of a coalition may, in general, depend also on the other players outside the coalition. However, we will first consider coalitional games in which the value of a coalition $\mathcal{S}$ depends only on the members of $\mathcal{S}$. Such coalitional games are said to be in characteristic form. Characteristic-form coalitional games are widely used, partly because of their simpler structure. Moreover, a coalitional game, regardless of whether it is in characteristic form or not, may have either transferable or nontransferable utility. For a coalitional game in characteristic form with transferable utility the value $v(\mathcal{S})$ of a coalition $\mathcal{S}$ is a real number. The value $v(\mathcal{S})$ quantifies the total reward/payoff of the coalition $\mathcal{S}$ that may be divided arbitrarily among its members. In spectrum sensing and access problems, the value of a coalition could, for example, be equivalent to the vacant bandwidth found by the SUs in the coalition that then could be shared among the coalition members using some fair rule. This fair rule would ensure that the effort a user contributes to the coalition would be rewarded through sharing the common resources. For example, if a user is able to find a lot of idle spectrum for the coalition, this may be taken into account when scheduling users. Alternatively, we could also define the value of a coalition as the sum throughput of its members. However, dividing the secondary system throughput arbitrarily among the coalition members is not generally straightforward since the throughput of each individual user depends on the channel quality between the corresponding receiver and transmitter pair as well as on the local interference.

Thus, the utility (throughput) cannot be arbitrarily transferred among the coalition members. Games having this latter characteristic are called *coalitional games with nontransferable utility.* In this case, the value of a coalition $\mathcal{S}$ is represented as a set of payoff vectors, $\upsilon(\mathcal{S}) \subseteq \mathbb{R}^{|\mathcal{S}|}$, that the members of $\mathcal{S}$ can achieve. Each element r_i of a payoff vector $r \in \upsilon(\mathcal{S})$ corresponds to a payoff member i can receive in coalition $\mathcal{S}$.

A coalitional game, in characteristic form, is superadditive if the value of a coalition formed by joining two disjoint coalitions is always at least equal to the value obtained by the two disjoint coalitions separately. Hence, cooperation is always beneficial, and a grand coalition, i.e., a coalition involving all the players, is the optimal coalition with the highest value. A coalitional game in characteristic form with superadditivity is called a *canonical coalitional game.* Canonical coalitional games focus on studying the stability of the grand coalition and finding a reward allocation that ensures that the players do not have an incentive to leave the grand coalition. A fundamental solution concept for canonical coalitional games is the core. The core of a coalitional game is a similar concept to the Nash equilibrium in noncooperative games. The core of a canonical coalitional game with transferable utility is a reward allocation $r = \{r_1, \ldots, r_N\}$ defined as

$$\mathcal{C} \triangleq \left\{ r : \sum_{i \in \mathcal{N}} r_i = \upsilon(\mathcal{N}) \text{ and } \sum_{i \in \mathcal{S}} r_i \geq \upsilon(\mathcal{S}), \forall \mathcal{S} \subseteq \mathcal{N} \right\}, \quad (11)$$

where $r_1, \ldots, r_N$ are the individual rewards of the players. Thus, no group of rational players has an incentive to leave the grand coalition whose reward allocation r is in the core of the game. The core of a coalitional game with transferable utility is obtained by solving the following linear program (if the core exists):

$$\min_{r} \sum_{i \in \mathcal{N}} r_i, \text{ s.t.} \sum_{i \in \mathcal{S}} r_i \geq \upsilon(\mathcal{S}), \forall \mathcal{S} \subseteq \mathcal{N}. \quad (12)$$

The core can be defined for other coalitional games as well.

Formulating a game such that the grand coalition is optimal may not always be the most appropriate model. In spectrum sensing and access problems, different costs and gains associated with cooperation combined with the local nature of the spectrum state may result in a situation in which the optimal coalition structure is not the grand coalition. Coalitional games in partition form are appropriate models for such scenarios. A coalitional game is in partition form if the value of a coalition $\mathcal{S}$ depends also on the players outside the coalition $\mathcal{S}$, i.e., players in $\mathcal{N} \setminus \mathcal{S}$, and how they are partitioned to other coalitions. That is, the value $\upsilon(\mathcal{S}, \Pi)$ of a coalition $\mathcal{S}$ depends on both the coalition $\mathcal{S}$ and the network partition Π.

A comprehensive treatment of coalitional games and their applications in communication networks can be found in [49].

COALITIONAL GAMES FOR JOINT
SPECTRUM SENSING AND ACCESS

In [50], coalitional games in partition form with nontransferable utility have been proposed for joint multiband spectrum sensing and access in cognitive radio ad hoc networks. SUs form coalitions to share local sensing statistics, coordinate the local sequential frequency band sensing orders, and cooperatively distribute their powers so that the total sum-rate of the coalition is maximized.

Each user maintains an ordered list of frequency bands in decreasing order of preference in which they would like to sense the frequency bands. In [50], the preference is modeled by weights $w_{i,k} = \theta_k g_{i,k}$, where θ_k is the probability that frequency band k is available and $g_{i,k}$ is the channel gain experienced by SU i on frequency band k. In a coalition, the individual frequency band sensing orders of the coalition members are then chosen cooperatively based on the local preference lists. The coalition members sort their lists starting from the highest-ranked frequency band and proceeding until the end as follows [50, Algorithm 1]: At each rank (i.e., list position), those SUs whose choices do not conflict with others are assigned these frequency bands. In case of a conflict, the user with the highest weight $w_{i,k}$ is assigned the frequency band. The remaining users repeat the procedure with the remaining frequency bands until each user has been assigned a frequency band for that particular rank in the list. If during this procedure a user is left without any possible frequency band in a current rank, the user selects the frequency band with the highest weight from frequency bands not already on its list (this results in interference with at least one other coalition member). This procedure is repeated for each rank of the ordered lists until each coalition member has a new cooperatively sorted sensing order.

Given the new sensing orders, the SUs then proceed to sense the frequency bands in the order of their lists until they find an available frequency band to access. The coalition members find available frequency bands simultaneously and then share their sensing results and cooperatively allocate their transmit powers on these frequency bands so that their total sum rate is maximized. The payoff of an SU i in a coalition $\mathcal{S}$ is modeled as [50]

$$r_i(\mathcal{S}, \Pi) = \bar{C}_i^{\mathcal{S}}(1 - \tau_i^{\mathcal{S}}), \quad (13)$$

where Π is the network partition, $\bar{C}_i^{\mathcal{S}}$ is the average capacity, which depends on channel gains, interference, sensing order, and frequency band availability, and $\tau_i^{\mathcal{S}}$ is the average fraction of time spent for sensing, which depends on the sensing order and frequency band availability. Therefore, the user payoffs depend on the network partition, which affects the coordination of sensing and access and, thus, the interference experienced by the users.

To maximize their payoffs, the SUs have to find an optimal network partition. In [50], this is addressed through a coalition formation algorithm in which potential coalition switches are periodically initiated in a random order by individual SUs. A switch from one coalition to another is approved only if the payoff of the switching SU is strictly increased without decreasing the payoff of any existing member of the new coalition. The users also maintain a history of their previous coalitions to ensure that they do not revisit any of the previous coalitions with smaller payoffs. The proposed algorithm is shown in [50] to converge to a Nash-stable partition in which no user has an incentive to switch to a different coalition. Furthermore, in practice, the proposed coalition formation algorithm can be run periodically with reset history tables.

This ensures discovery of new neighbors and adaptation to environmental changes.

The simulation results in [50] show that the proposed coalitional games result in a significantly higher average payoff than noncooperative sensing and access. The simulation results also show that the proposed coalition formation game is likely to converge to network partitions consisting of many small coalitions. This indicates that the grand coalition is not the optimal coalition and increasing the coalition size too much decreases the chances of exploiting the best frequency bands due to a less favorable frequency band ordering within the coalition.

REPEATED GAMES
All of the games we have considered so far are single-stage games, i.e., the games are played for only one round. A repeated game is a game in which the same-stage game is repeated multiple times. In repeated games, the players have to take into account the impact of their current actions on the future rewards. The most interesting form of repeated games for spectrum sensing and access problems is an infinitely repeated game. In infinitely repeated games, the players' objective functions can be, e.g., discounted sums of rewards.

STOCHASTIC GAMES
Although infinitely repeated games can be used to model continuous multiuser spectrum sensing and access problems, they still lack the dynamic nature common to these problems. In cognitive radio applications, the state of the spectrum and the surrounding radio environment is dynamically changing as a function of time and location. In repeated games, and all of the other games described in this article, the players are faced with the same-stage game each time. Hence, the players' strategies do not depend on the state of the environment. Stochastic games can be used to model the dynamic environments faced in the most general spectrum sensing and access problems. In stochastic games, the environment has multiple states that change stochastically during the course of the game. Thus, the players' strategies depend on the current state.

Stochastic games are extensions of MDPs to the case of multiple players. A stochastic game consists of

- a sequence of discrete time steps $n = 0, 1, 2, \ldots$
- a set of players $\mathcal{N} = \{1, \ldots, N\}$
- a set of possible states $s \in S$
- a set of possible actions for each player in each state $a_i \in A_i, i \in \mathcal{N}$; the combined action space is $A = A_1 \times \cdots \times A_N$
- a state transition function $\varphi : S \times A \times S \to [0, 1]$ that defines the state transition probability $P(s_{n+1} \mid s_n, a_{1,n}, \ldots, a_{N,n})$
- reward (payoff) functions $r_i : S \times A \times S \to \mathbb{R}, i \in \mathcal{N}$ that give the players rewards $r_{i,n+1}$ for the joint action $a_{1,n}, \ldots, a_{N,n}$ of the players in state s_n resulting in new state s_{n+1}.

Similarly to POMDPs, stochastic games can be generalized to partially observable stochastic games [16] in which the state is observable only through a stochastic function. Stochastic games form the basis for modeling multiagent reinforcement learning problems.

STOCHASTIC GAMES FOR SPECTRUM SENSING AND ACCESS
In [14], the multiuser and multiband spectrum sensing and access problem is formulated as a noncooperative partially observable stochastic game in which the SUs compete for spectrum opportunities. Each SU is able to sense only one frequency band in each time slot, which results in partial observability of the PU state. The existence of a symmetric Nash equilibrium in which all SUs play the same strategy is established. In addition, a Stackelberg game is proposed for improving the Nash equilibrium. A Stackelberg game is a game consisting of a leader (network manager), who plays first, and followers (SUs), who play second. In [14], the network manager aims to maximize the average total throughput by controlling the apparent availability of the PU frequency bands. That is, the network manager can reserve a PU frequency band which then appears to the SUs as being occupied. This results in decreased competition among the SUs, which then increases the average total throughput.

DYNAMIC PROGRAMMING
Dynamic programming refers to an optimization approach in which an original multistage sequential decision problem is broken down into smaller, simpler single-stage decision problems that are then solved in a recursive manner. Here, we consider solving finite MDPs using dynamic programming. Therefore, dynamic programming can be used to solve spectrum sensing and access problems formulated as finite MDPs.

We will consider the maximization of (9), i.e., the expected sum of discounted rewards given an initial state $s_n = s$. The value of a policy π starting from state s is defined by

$$\mathcal{V}^\pi(s) = E_\pi \left[\sum_{k=0}^{\infty} \gamma^k r_{n+k+1} \mid s_n = s \right], \qquad (14)$$

where $E_\pi[\cdot]$ denotes the expectation when the policy π is followed. The value function can be broken down as follows:

$$\begin{aligned} \mathcal{V}^\pi(s) &= E_\pi \left[\sum_{k=0}^{\infty} \gamma^k r_{n+k+1} \mid s_n = s \right] \\ &= E_\pi \left[r_{n+1} + \gamma \sum_{k=0}^{\infty} \gamma^k r_{(n+1)+k+1} \mid s_n = s \right] \\ &= E_\pi \left[r_{n+1} + \gamma \mathcal{V}^\pi(s_{n+1}) \mid s_n = s \right], \end{aligned} \qquad (15)$$

where the last form in (15) is called the *Bellman equation* for $\mathcal{V}^\pi(s)$ [56]. The optimal value function $\mathcal{V}^*$ is equal to the expected return of the best action in state $s_n = s$

$$\mathcal{V}^*(s) = \max_a E \left[r_{n+1} + \gamma \mathcal{V}^*(s_{n+1}) \mid s_n = s, a_n = a \right]. \qquad (16)$$

This is called the *Bellman optimality equation* [56].

From (15) and (16), we see that successive states have a recursive relationship. Moreover, the optimal policy from any state s_{n+1} does not depend on what happened before that state was reached. Thus, the original problem can be solved recursively using the Bellman optimality equation. In finite-horizon problems, we can start from the last time step and work recursively

backward using the Bellman optimality equation. This is called *backward induction*. Figure 5 illustrates backward induction in a problem of finding the longest path in a network of nodes.

In infinite horizon problems, the main solution approaches are value iteration and policy iteration. Value iteration is a generalization of backward induction in which the Bellman optimality equation is used iteratively to update the value function until it converges. Policy iteration starts from an arbitrary policy that is then successively evaluated and improved through greedy action selection in each state until it converges. Each successive policy is guaranteed to be strictly better than the previous policy until the optimum is found. For more information on value and policy iteration, see [56].

Dynamic programming is computationally more efficient than evaluating all possible policies. Nevertheless, the computational and memory requirements of dynamic programming grow exponentially as the number of states and actions increases, i.e., dynamic programming suffers from the curse of dimensionality. Moreover, in general, the state transition probabilities and rewards have to be known to evaluate the Bellman equations. Thus, the applicability of dynamic programming is mostly limited to problems involving only small state and action spaces with known state transition probabilities and rewards. However, dynamic programming provides insight into the optimal solution that can then be employed to obtain practical near-optimal reduced-complexity algorithms and learning methods.

DYNAMIC PROGRAMMING-BASED OPTIMIZATION OF THE SENSING ORDER OVER MULTIPLE FREQUENCY BANDS

In [19], a dynamic programming solution has been proposed for the problem of selecting the sequential order of sensing different frequency bands in a multiband single-user cognitive radio system. The goal is to maximize the throughput of the SU that senses frequency bands in a sequential order until it finds a frequency band that is both vacant and has acceptable channel quality. In [19], this problem is formulated as a sequential decision problem and dynamic programming is employed to solve it. The solution is obtained with backward induction similar to the longest path example in Figure 5. Thus, the optimal sensing order is found by working backward from the end of the sensing order and selecting at each stage the sensing order for the remaining frequency bands that maximizes the expected throughput. In this case, the states correspond to the sets of frequency bands already sensed. Scenarios with both known and unknown frequency band availability probabilities are considered.

BANDIT PROBLEMS

Bandit problems are sequential decision problems in which the name originates from the similarity to the traditional slot machines used in casinos, called *one-armed bandits*. Bandit problems can be modeled as MDPs. Hence, bandit problems are appropriate models, in particular, for single-user multiband spectrum sensing and access problems. That is, multiband spectrum sensing and access problems can be modeled as multiarmed bandit problems in which the arms of the multiarmed bandit correspond
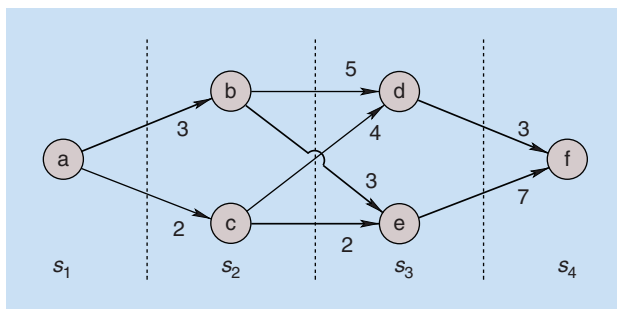
[FIG5] A dynamic programming example: finding the longest path from *a* to *f*. The thicker arrows indicate the longest path from each state to *f*. The longest path from *a* to *f*, i.e., $a \rightarrow b \rightarrow e \rightarrow f$, has length 13. This optimal path can be found recursively starting from *f*: Setting $\mathcal{V}(f) = 0$, it follows from the Bellman optimality equation (with $\gamma = 1$) that $\mathcal{V}(d) = r(d, f) + \mathcal{V}(f) = 3 + 0 = 3$ and $\mathcal{V}(e) = r(e, f) + \mathcal{V}(f) = 7 + 0 = 7$. Now, the Bellman optimality equation for $s_2 = b$ gives $\mathcal{V}(b) = \max\{r(b, d) + \mathcal{V}(d), r(b, e) + \mathcal{V}(e)\} = \max\{5 + 3, 3 + 7\} = 10$. Similarly, one can calculate $\mathcal{V}(c) = 9$. The final result $\mathcal{V}(a) = 13$ follows again from another recursive application of the Bellman optimality equation.

to different frequency bands and choosing an arm to play corresponds to sensing or accessing a particular frequency band.

In the following, we will consider Markovian multiarmed bandit problems in which the conditional state transition probability depends only on the current state, and, thus, they can be modeled as MDPs. There are K arms and a decision maker controlling the selection of the arms to play. In each state s_n, the decision maker selects one arm a_i to play and receives a reward r_{n+1} for it. In classical multiarmed bandit problems, the state of the nonplayed arms does not change and the stochastic processes for the different arms are independent of each other.

Gittins and Jones [13] showed that the optimal policy maximizing the expected sum of discounted rewards with known rewards and state transition probabilities is given by an index structure in which a priority index, called the *Gittins index,* is calculated for each state of each arm of the multiarmed bandit. The Gittins index of each arm depends only on that arm's underlying stochastic process. Thus, the original K-dimensional optimization problem is reduced to K one-dimensional optimization problems. Once the Gittins indices have been calculated, the optimal policy reduces to selecting the arm with the largest Gittins index at each time. The Gittins index for the kth arm is given by

$$v_k(s_0^k) = \max_{\tau > 0} \frac{E\left[\sum_{n=0}^{\tau-1} \gamma^n r_{n+1}^k(s_n^k) \mid s_0^k\right]}{E\left[\sum_{n=0}^{\tau-1} \gamma^n \mid s_0^k\right]}, \qquad (17)$$

where the maximization is over the set of all stopping times $\tau > 0$. Here, $r_{n+1}^k(s_n^k)$ denotes that the reward of the kth arm r_{n+1}^k depends only on the state of the kth arm s_n^k. Moreover, if an arm is not played its reward is zero. The Gittins index, thus, finds for each arm the optimal stopping time in terms of maximizing the expected discounted reward normalized by the expected discounted number of plays. Several algorithms have been proposed for efficient calculation of Gittins indices; see, e.g., [39] and the references therein. Note, however, that calculating the Gittins
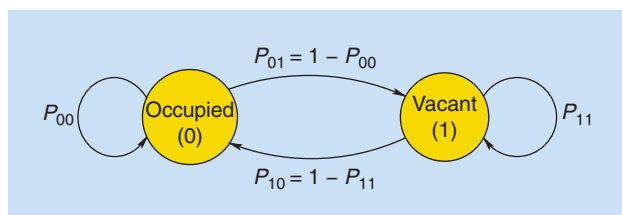
[FIG6] A two-state Markov chain model for frequency band occupancy.

indices requires, in general, full knowledge of the rewards and state transition probabilities for the arms.

The classical multiarmed bandit problem has a few limitations that make it not the most appropriate model for multiband spectrum sensing and access problems. First, and most importantly, the state of an arm can change only when it is played. Second, only one arm can be played at a time. In multiband spectrum sensing and access problems, the state of a frequency band depends on the PU activity and, hence, it can change at any time regardless of SU actions. Furthermore, the SUs may have the capability to sense or access multiple frequency bands simultaneously depending on the transceiver front-end properties. In the following, we focus on restless multiarmed bandit problems that remove both of these limitations.

RESTLESS MULTIARMED BANDIT PROBLEMS
The restless multiarmed bandit problem is a generalization of the classical multiarmed bandit problem in which the decision maker may simultaneously play multiple $L \leq K$ arms and the nonplayed arms may change state and give rewards [63]. This complicates the problem significantly and renders the Gittins index policy suboptimal. The problem is further complicated if the arms are dependent. In the following, we will, however, focus on restless multiarmed bandit problems in which the arms are independent. For this problem when the rewards and state transition probabilities are known, Whittle proposed in [63] an index policy based on a Lagrange multiplier approach that is optimal for the average expected reward over the infinite horizon criterion under a relaxed constraint that the average number of played arms is equal to L. However, calculating the Whittle indices may be very difficult in practice. Moreover, the problem may not even be indexable, i.e., the ordering of the arms given by the Whittle index may not be consistent and thus meaningful. In the following, we will illustrate through an example [29] how the single-user multiband spectrum sensing and access problem can be formulated as a restless multiarmed bandit problem.

WHITTLE INDEX POLICY FOR SINGLE-USER
MULTIBAND SPECTRUM SENSING
In [29], a single-user multiband spectrum sensing problem is formulated as a restless multiarmed bandit problem with N_B arms each corresponding to a single PU frequency band. The SU senses $K \leq N_B$ frequency bands in each time slot (either sequentially or simultaneously). For the vacant frequency bands, the user receives a reward equal to the transmission rate r_i, $i = 1, \ldots, N_B$, that is in general different for every frequency band. For the

other frequency bands, the reward is zero. Sensing is assumed to be error free. The goal of the SU is to maximize the reward over the infinite time horizon; two performance criteria are considered in [29]: the expected discounted reward over the infinite horizon and the expected average reward over the infinite horizon. The PU occupancies on the frequency bands are modeled using independent Markov chains with two possible states [vacant (1) and occupied (0)] and known state transition probabilities (P_{01}^i, P_{11}^i), $i = 1, \ldots, N_B$; see Figure 6. The user observes the frequency band states only after sensing and, hence, needs to infer the state from its past decisions and observations to make decisions. The conditional probability that a frequency band is in state 1 given all past decisions and observations is a sufficient statistic [55]. The vector of conditional probabilities is referred to as the *belief vector* $\Omega_n = [\omega_n^1, \ldots, \omega_n^{N_B}]$. The belief ω_n^i that the frequency band i, $i = 1, \ldots, N_B$, is vacant can be updated recursively [29]

$$\omega_{n+1}^i = \begin{cases} P_{11}^i, & i \text{ sensed vacant,} \\ P_{01}^i, & i \text{ sensed occupied,} \\ \omega_n^i P_{11}^i + (1 - \omega_n^i) P_{01}^i, & i \text{ not sensed.} \end{cases} \quad (18)$$

Consequently, the state of the ith arm at time n is given by the belief state ω_n^i. This restless multiarmed bandit problem is shown to be indexable in [29]. Moreover, the Whittle index has been obtained in [29] in closed form for both of the reward criteria.

This problem can also be viewed as a POMDP with independent frequency bands. A more general POMDP formulation of single-user spectrum sensing and access with possibly correlated frequency bands can be found in [12] and [65] where an optimal solution with known state transition probabilities is also derived. In addition, a scenario with nonideal sensing is considered. In both cases, the computational complexity of the optimal solution grows exponentially with the number of frequency bands [65].

In general, calculating the Whittle index requires full knowledge of the transition probabilities P_{11}^i and P_{01}^i as well as the rewards r_i. However, for i.i.d. arms (i.e., all arms have equal transition probabilities and rewards), the Whittle index policy has been shown in [29] to be equivalent to a myopic (greedy) policy.

MYOPIC SPECTRUM SENSING POLICIES
A myopic policy chooses at each time the action maximizing the expected immediate reward while fully ignoring the impact on any future rewards [1], [30], [66]. Thus, there is no exploration. A myopic policy is always exploiting the action that gives the highest expected reward. The myopic action $\hat{a}_n = [\hat{a}_n^1, \ldots, \hat{a}_n^{N_B}]$, $\hat{a}_n^i \in \{0, 1\}$ for sensing M frequency bands at time n, is given by

$$\hat{a}_n = \arg \max_{a_n} \sum_{i=1, a_n^i \neq 0}^{N_B} \omega_n^i r_i, \quad \text{s.t.} \sum_{i=1}^{N_B} a_n^i = M, \quad (19)$$

where ω_n^i is the belief state of frequency band i, $a_n^i = 1$ denotes that frequency band i is sensed, and $a_n^i = 0$ does the opposite. If the arms are i.i.d., the myopic policy admits a queue structure that depends only on the ordering of P_{11} and P_{01} [30], [66]. The ordering of the frequency bands is maintained with a queue, and, at each
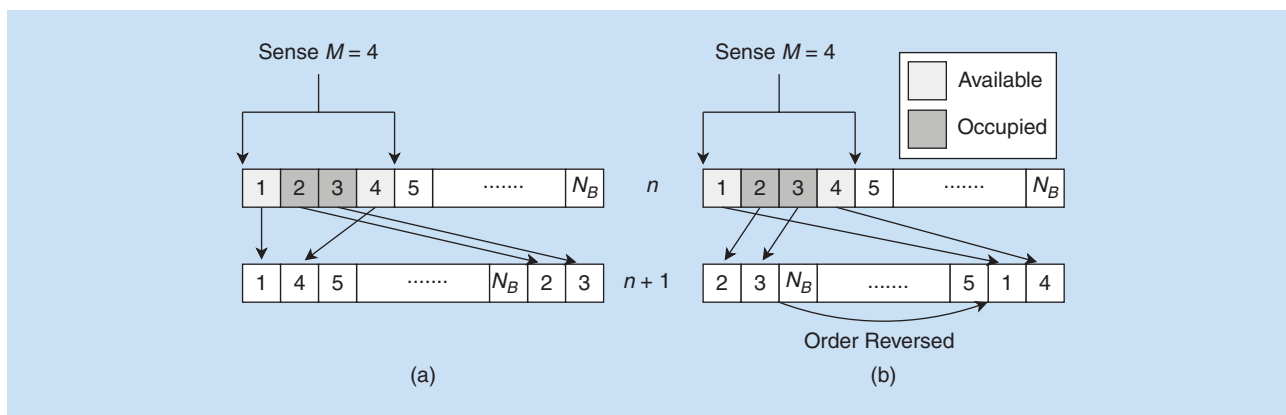
[FIG7] The structure of the myopic sensing policy for i.i.d. frequency bands when (a) $P_{11} \geq P_{01}$ and (b) $P_{11} < P_{01}$ [30].

time, the M frequency bands at the head of the queue are sensed. Hence, the belief states do not need to be updated, and the exact values of state transition probabilities are not needed. Only the ordering of P_{11} and P_{01} has to be known. This makes the myopic policy computationally very efficient to employ in practice. Figure 7 depicts the queue structure of the myopic policy.

For positively correlated ($P_{11} \geq P_{01}$) i.i.d. frequency bands, the myopic sensing policy is optimal for any M [1], [66]. This optimality holds for discounted expected reward over finite and infinite horizons and for average expected reward over the infinite horizon [1]. For negatively correlated ($P_{11} < P_{01}$) i.i.d. frequency bands, the myopic sensing policy has been shown in [1] to be optimal for $N_B = 2$ and $N_B = 3$ and $M = 1$ but, in general, not optimal for $N_B > 4$.

All of the results are under the assumption of error-free sensing. The myopic policy with imperfect sensing has been considered in [30]. In this case, the myopic policy follows the same queue structure as in the case of perfect sensing under a certain condition on the false alarm probability. Moreover, it remains optimal at least for $N_B = 2$ [30]. The performance of the myopic policy in [66] is evaluated in the next section (Figure 8).

UPPER CONFIDENCE BOUND ALGORITHMS FOR SPECTRUM SENSING AND ACCESS

A branch of bandit problem research has focused on deriving index policies based on upper confidence bounds for multiarmed bandits with unknown state transition probabilities and rewards [2], [8], [40]. Most of this work assumes that the rewards for each
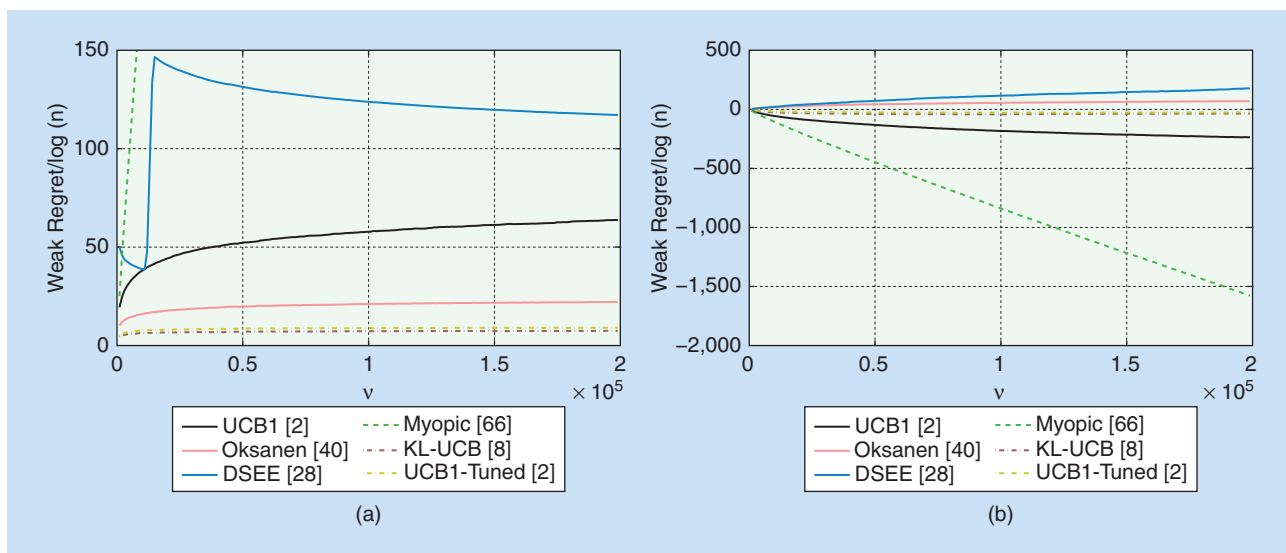
[FIG8] Normalized median weak regret averaged over more than 1,000 randomly chosen Monte Carlo experiments for different sensing policies. The number of frequency bands is ten with each frequency band producing different randomly chosen (a) i.i.d. Bernoulli ($P_{11} \sim \text{Uniform}(0, 1)$ and $P_{00} = 1 - P_{11}$) and (b) Markovian [$P_{11} \sim \text{Uniform}(0.7, 0.75)$ and $P_{00} \sim \text{Uniform}(0.65, 0.7)$] rewards. Upper confidence bound algorithms have good performance in scenarios in which it is likely that one frequency band is clearly better than the rest (a). In (b), the frequency bands are close to identical, and, hence, any single-arm policy sought by the upper confidence bound algorithms is substantially worse than the myopic policy that is optimal for i.i.d. frequency bands with Markovian rewards. However, the myopic policy fails completely when there is a significant disparity among the frequency band statistics as in (a).

arm are i.i.d. over time, albeit in general different for each arm. However, the ideas and the corresponding policies can be used also with other types of reward distributions such as Markovian rewards. The goal of these policies is to find the best arm producing the highest average rewards. In these policies, the index for each arm consists of the average reward and a confidence term. The confidence term grows as a function of time and reduces when the arm is played. For example, in [40], a single-user spectrum sensing policy was proposed that proceeds as follows:

1) Sense each frequency band, $i = 1, \ldots, N_B$, once. Thus, $n = N_B$.

2) For $n > N_B$ sense, the band with the highest index

$$I_{i,n} = \bar{r}_{i,n} + \sqrt{\ln\left(\frac{n}{n_i}\right)}, \quad i = 1, \ldots, N_B, \qquad (20)$$

where $\bar{r}_{i,n}$ is the average obtained reward for frequency band i (the rewards $r_{i,n}$ are assumed to be bounded in $[0, 1]$, $\forall i, n$) and n_i is the last time instant when band i was sensed.

We can observe from (20) that the role of the second term on the right-hand side is to promote exploration of suboptimal frequency bands at least occasionally and, thus, obtain a desired balance between exploitation and exploration. In this case, the time difference between two sensing instances of a suboptimal frequency band grows exponentially in time. This also means that the policy asymptotically achieves a logarithmic order of weak regret [40]. Asymptotically logarithmic order of regret is the optimum that can be achieved for i.i.d. rewards [24]. Note, however, that in weak regret the loss in performance is not compared to the globally optimal policy but to the optimal single-arm policy that plays only the arm with the highest expected reward. The optimal single-arm policy is the globally optimal policy for i.i.d. rewards. Another reason weak regret is commonly employed is that the optimal policy for weak regret is also much easier to obtain in the general case.

Figure 8 shows a performance comparison of several different upper confidence bound algorithms and the myopic policy [66] for multiband single-user spectrum sensing. The results show that the upper confidence bound algorithms are good for scenarios in which there is one dominating frequency band, while in scenarios with i.i.d. frequency bands with Markovian rewards, other policies, such as the myopic policy [66], are more suitable.

DECENTRALIZED MULTIUSER SENSING AND ACCESS POLICIES BASED ON BANDIT PROBLEMS
Although bandit problems are best suited for modeling and solving multiband single-user sensing and access problems, they can also be used in multiuser sensing and access problems. In [28], decentralized multiuser and multiband sensing and access policies have been proposed based on upper confidence bounds as well as on deterministic sequencing of exploration and exploitation. The SUs operate independently without constant information exchange but possibly with

> **STOCHASTIC GAMES CAN BE USED TO MODEL THE DYNAMIC ENVIRONMENTS FACED IN THE MOST GENERAL SPECTRUM SENSING AND ACCESS PROBLEMS.**

some preagreement about sharing the frequency bands. The requirement for preagreement can be relaxed with randomization during exploitation that results in a bounded loss in performance [28]. Thus, the proposed policy achieves logarithmic order of weak regret even without preagreement [28]. Collisions are assumed to occur when multiple SUs try to sense and access the same frequency band at the same time. Moreover, the users are assumed to be able to detect collisions, which, in practice, is difficult to achieve.

REINFORCEMENT LEARNING
Reinforcement learning is a trial-and-error machine-learning approach in which the decision maker, called the *agent,* observes the state of the environment and chooses actions that lead to rewards and new states. Actions leading to desired outcomes are given higher rewards, which reinforce these actions, thus making them more likely to be chosen again in similar situations in the future. Consequently, in reinforcement learning, the agent or agents are faced with the exploitation versus exploration tradeoff, i.e., whether to exploit the current best action or to explore other actions in hope of finding a better one. Single-user reinforcement learning problems can be modeled as MDPs while multiagent reinforcement learning problems can be modeled using stochastic games. Bandit problems comprise one of the simplest reinforcement learning problems.

SINGLE-AGENT REINFORCEMENT LEARNING
The goal of single-agent reinforcement learning is for the agent to learn a policy that optimizes the cost function, such as the expected sum of discounted rewards (9). One way of achieving this is to learn an optimal action-value function. The action-value function $\mathcal{Q}(s, a)$ evaluates the value of each action in a given state. It is defined as the expected return of taking an action a in state $s_n = s$ and then following policy π, i.e.,

$$\mathcal{Q}^\pi(s, a) = E_\pi\left[\sum_{k=0}^{\infty} \gamma^k r_{n+k+1} \,\big|\, s_n = s, a_n = a\right]. \qquad (21)$$

The optimal action-value function $\mathcal{Q}^*$ satisfies the Bellman optimality equation [56, p. 76]

$$\begin{aligned} \mathcal{Q}^*(s, a) &= \max_\pi \mathcal{Q}^\pi(s, a) \\ &= \sum_{s' \in S} P(s' \mid s, a)[r(s, a, s') + \gamma \max_{a'} \mathcal{Q}^*(s', a')], \end{aligned} \qquad (22)$$

where $r(s, a, s')$ is the reward function. Once the agent has computed the optimal action-value function $\mathcal{Q}^*$, it can employ the greedy policy, which always chooses the action maximizing $\mathcal{Q}^*$ in the current state, i.e., $\hat{a} = \arg\max_{a \in A} \mathcal{Q}^*(s, a)$, to achieve its goal.

One of the key concepts in reinforcement learning is temporal-difference learning, and $\mathcal{Q}$-learning [62] is its most celebrated algorithm. $\mathcal{Q}$-learning is a model-free off-policy temporal-difference learning algorithm. The action-value updates of $\mathcal{Q}$-learning for each state-action pair are given by

$$\mathcal{Q}_{n+1}(s_n, a_n) = \mathcal{Q}_n(s_n, a_n)$$
$$+ \alpha_n \left[r_{n+1} + \gamma \max_{a' \in A} \mathcal{Q}_n(s_{n+1}, a') \right.$$
$$\left. - \mathcal{Q}_n(s_n, a_n) \right], \qquad (23)$$

where α_n $(0 < \alpha_n \leq 1)$ is a step size parameter (learning rate). The action-value function update of $\mathcal{Q}$-learning is given by the temporal difference between the current estimate $\mathcal{Q}_n(s_n, a_n)$ and the target value $r_{n+1} + \gamma \max_{a' \in A} \mathcal{Q}_n(s_{n+1}, a')$ multiplied by the step size α_n. Hence, knowledge of the state transition probabilities and the reward function is not required. The action-value function of $\mathcal{Q}$-learning approximates the optimal action-value function $\mathcal{Q}^*$ regardless of the followed policy. Thus, $\mathcal{Q}$-learning is an off-policy temporal-difference algorithm.

In $\mathcal{Q}$-learning, the policy followed can essentially be any policy as long as all state-action pairs are visited infinitely many times. Opposite to this are on-policy algorithms in which the action-value function updates are based on the policy followed by the agent and the state-action pairs visited. Sarsa [48] is an example of an on-policy temporal-difference algorithm. The action-value updates of the one-step Sarsa algorithm are defined by

$$\mathcal{Q}_{n+1}(s_n, a_n) = \mathcal{Q}_n(s_n, a_n)$$
$$+ \alpha_n [r_{n+1} + \gamma \mathcal{Q}_n(s_{n+1}, a_{n+1})$$
$$- \mathcal{Q}_n(s_n, a_n)]. \qquad (24)$$

The update is given by the temporal difference between the $\mathcal{Q}$-values of two consecutive state-action pairs visited by the algorithm at times n and $n+1$ multiplied by α_n.

The convergence of the action-value function $\mathcal{Q}_n$ to the optimum $\mathcal{Q}^*$ in stationary environments has been established for $\mathcal{Q}$-learning in [62] and for one-step Sarsa in [54]. The convergence to $\mathcal{Q}^*$ is guaranteed with probability one if the agent employs a lookup-table to store the $\mathcal{Q}$-values for every state-action pair, visits every state-action pair infinitely many times, and the step size parameter is chosen such that $\sum_{n=0}^{\infty} \alpha_n = \infty$ and $\sum_{n=0}^{\infty} \alpha_n^2 < \infty$. Furthermore, to ensure the convergence of one-step Sarsa, the learning policy must become greedy in the limit.

Thus, both $\mathcal{Q}$-learning and Sarsa require a learning policy that balances between exploitation and exploration. The ϵ-greedy action selection is one of the most commonly employed policies in $\mathcal{Q}$-learning and Sarsa. It is a simple method that balances between exploitation and exploration by selecting the action that maximizes the action-value function, i.e., $a^* = \arg \max_a \mathcal{Q}(s, a)$, with a probability $1 - \epsilon$, or a random action, uniformly, with probability ϵ regardless of the action-value function estimates. Another commonly employed simple action selection method is the softmax action selection method. The softmax method chooses action a in state s with probability

> **A KEY ASSUMPTION THAT FACILITATES FINDING A TRACTABLE SOLUTION TO THE DISTRIBUTED DETECTION PROBLEM IS THE CONDITIONAL INDEPENDENCE OF OBSERVED SENSOR DATA CONDITIONED ON THE HYPOTHESIS.**

$$\pi(s, a) = \frac{\exp(\mathcal{Q}(s, a)/\tau)}{\sum_{a' \in A} \exp(\mathcal{Q}(s, a')/\tau)}, \qquad (25)$$

where $a \in A$ and τ is a positive temperature parameter controlling the weighting of different actions. Low temperatures increase the differences in the action selection probabilities, while high temperatures cause all actions to be almost equiprobable. In the limit, when $\tau \to 0$, the softmax method corresponds to greedy action selection, and when $\tau \to \infty$, the actions are selected randomly from a uniform distribution.

In practice, the learning environment in cognitive radio applications is usually nonstationary. Thus, the goal is to track a nonstationary state of the spectrum. A practical approach is to use a constant step size $\alpha_n = \alpha, \forall n$. A constant step size does not satisfy the previously described conditions for convergence. However, it guarantees that the more recent samples have larger weights than the ones in the distant past, which facilitates tracking the solution of a nonstationary problem. In addition, it is important to continue to explore in nonstationary environments. This can be accomplished, for example, by employing the ϵ-greedy algorithm with a constant nonzero ϵ.

MULTIAGENT REINFORCEMENT LEARNING

A multiagent reinforcement learning problem can be formulated as a stochastic game. In reinforcement learning, the players of the stochastic game are called *agents,* as in the single agent case. A multiagent reinforcement learning problem is much more complicated than a single-agent reinforcement learning problem. In a multiagent reinforcement learning problem, the rewards and state transitions depend, in general, on the joint actions of all the agents. Thus, the objective function, e.g., the expected sum of discounted rewards in (9), depends also on the joint actions of all the agents. Hence, the agents have to adapt their own policies not only to the environment but to the other agents' policies as well.

Multiagent reinforcement learning problems can be either noncooperative or cooperative. If the reward functions of the competing agents sum to zero, the problem is fully competitive. If the interests of the different agents are fully aligned, the reward functions are identical for all agents and the problem is fully cooperative. Problems without any restrictions on the reward functions of the agents are the most general ones and can have elements of both competition and cooperation. Most of the existing multiagent reinforcement learning algorithms are based on $\mathcal{Q}$-learning [7]. Different algorithms include different levels of coordination in both learning algorithm and action selection. In cooperative problems, the level of coordination among the agents has a significant effect on both the stability of the learning process and the adaptation to the other agents' policies. The level of coordination depends also on the application and its restrictions. In some tasks, there may be explicit coordination in the form of pre-established preferences toward certain joint actions, or actions may, e.g., be selected in turn and

communicated to other agents. In the absence of any explicit coordination mechanism, the agents may try to learn a model of the other agents' policies provided that they are able to observe the other agents' actions. For a comprehensive review of different multiagent reinforcement learning algorithms and their characteristics, see [7].

MULTIAGENT REINFORCEMENT LEARNING FOR DISTRIBUTED MULTIBAND SPECTRUM SENSING

In [34], the multiuser multiband spectrum sensing problem is formulated as a partially observable stochastic game among the SUs. The proposed game consists of a group of SUs, a set of possible PU states (idle or occupied for each frequency band), a set of possible actions (which band to sense for each SU), observations corresponding to sensing decisions for the sensed frequency bands, a state transition function, reward functions (the number of frequency bands sensed vacant), and observation (sensing decision) probability functions. In the proposed formulation, the SUs cooperate to maximize the expected amount of (discounted) vacant spectrum found locally at each SU under a constraint on the probability of PU detection. The SUs may sense only a subset of the frequency bands in each sensing period. Moreover, the SUs employ local detection algorithms subject to decision errors. These two limitations make the PU state only partially observable.

The cooperation among the SUs is achieved through local interaction in [34]. The SUs exchange information for two reasons. First, the SUs exchange their local sensing statistics to enable collaborative distributed detection. Second, the SUs exchange the indices of the frequency bands they are going to sense in the next sensing period to coordinate their actions.

The proposed cooperative multiagent reinforcement learning approach is based on Sarsa with linear function approximation. In linear function approximation, the action-value function Q is

> **STOCHASTIC GAMES FORM THE BASIS FOR MODELING MULTIAGENT REINFORCEMENT LEARNING PROBLEMS.**

approximated with a linear function. In [34], the linear function is given by

$$Q_n^k(\omega_n^k, a_n^k) = (\theta_n^k)^T f(\omega_n^k, a_n^k) = \sum_{i=1}^{N_B} \theta_{n,i}^k f_i(\omega_n^k, a_n^k), \qquad (26)$$

where θ_n^k is a parameter vector and $f(\omega_n^k, a_n^k)$ is a feature vector depending on the belief state ω_n^k and the actions of SU k and its neighbors. The belief state ω_n^k indicates SU k's belief that each frequency band is vacant. The actions a_n^k correspond to the indices of the frequency bands sensed by SU k and its neighbors. The belief state update follows a similar form as (18) but is more complicated since it also takes into account the probabilities of detection errors and involves learning the state transition probabilities. The feature value depends on both the SU's belief that the spectrum is vacant and the probability of a false alarm, and is given by [34]

$$f_i(\omega_n^k, a_n^k) = \omega_n^{k,i} \cdot h\left(I_{[a_n^k = i]} + \sum_{j \in G_k} I_{[a_n^j = i]} \right), \qquad (27)$$

where $h(m) = 1 - p_{fa,m}$, with $p_{fa,m}$ the false alarm probability obtained with m SUs sensing, and $I_{[a_t = i]}$ is an indicator function having value 1 if $a_n = i$ and 0 otherwise. Thus, the argument of the function h is the number of SUs in group $k \cup G_k$ sensing the frequency band i in the time slot n where G_k denotes SU k's neighbors.

The obtainable false alarm probability depends on the propagation environment and the local sensing algorithm. In difficult propagation environments, a larger number of SUs is required to achieve a low probability of false alarm. As a result, the SUs also try to learn the optimal level of sensing cooperation for each frequency band in each particular location so that the expected amount of idle spectrum found is maximized.

Approximating the action-value function Q with a linear function transforms the problem of learning the action-value function Q to the problem of learning the parameter vector θ_n. In [34], the learning algorithm is gradient-descent-based Sarsa with ϵ-greedy action selection.

Figure 9 shows the spectrum sensing performance of the learning approach proposed in [34] in a cognitive radio scenario with ten cooperating SUs and seven PU frequency bands in both AWGN and Rayleigh fading PU to SU channels.

[FIG9] The average sensing performance over 1,000 Monte Carlo runs of the algorithm proposed in [34] relative to an optimal genie in AWGN and Rayleigh fading PU to SU channels with ten SUs and seven PUs. In the Rayleigh fading propagation environment, the learning of the state transition probabilities is slower because of lower sensing reliability. This also affects the overall sensing performance in the beginning of the simulation.

OTHER REINFORCEMENT LEARNING-BASED SENSING AND ACCESS POLICIES

In [42], a single-agent reinforcement learning-based cooperative multiuser and multiband sensing policy is proposed. The proposed sensing policy consists of two stages both coordinated by an FC. In the first stage, the goal of the FC is to choose the frequency bands with the highest expected throughput for sensing to maximize the throughput of the cognitive radio network. In

the second stage, the FC assigns the SUs to sense the frequency bands selected in the first stage such that the probabilities of missed detection are minimized on the chosen frequency bands. The learning approach employed in [42] uses single-state $\mathcal{Q}$-learning with ϵ-greedy action selection.

In [32], a cooperative single-band sensing approach based on single-agent temporal-difference learning is proposed. The goal of the FC controlling the sensing policy is to find an optimal subset of SUs to perform sensing in an optimal sequential order so that the reporting delay and overhead are minimized.

CONCLUSIONS AND FUTURE DIRECTIONS

In this article, we have provided an overview of spectrum exploration and exploitation methods for cognitive radio systems. In the first part of the article, we provided a brief introduction to advanced spectrum sensing techniques, such as distributed detection, SD, and quickest detection. Distributed detection is important for mitigating propagation effects and, thus, improving the reliability of spectrum sensing. Sequential and quickest detection techniques aim at minimizing the time spent in sensing a particular frequency band and, hence, facilitate using more time for transmissions and increase the throughput. In the second part of the article, we presented various different approaches for spectrum sensing and access policy design in cognitive radio networks. We have shown how to formulate the spectrum sensing and access problems as MDPs or using game theory. We have provided brief introductions to dynamic programming, bandit problems, reinforcement learning, and game theory, and have reviewed the various state-of-the-art spectrum sensing and access policies based on these techniques.

This tutorial article shows that considerable advancement has been achieved in recent years in the field of flexible spectrum use. However, further progress is still needed to fully realize the goal of secondary opportunistic spectrum use and efficient spectrum exploration and exploitation. We expect the following three important design aspects to play a major role in the future development of joint spectrum sensing and access algorithms and methods:

1) *Dynamic problem formulation*: The radio-frequency spectrum is a time-varying resource. Thus, dynamic game-theoretic models, such as stochastic games, will play an important role in modeling realistic multiuser spectrum sensing and access problems.

2) *Partial observability and other limitations*: In practical cognitive radio systems, the SUs operate with limited resources. This results in partial observability of the spectrum state. Furthermore, the observations are subject to errors. Taking these aspects into account and analyzing how they will affect the performance of learning algorithms and various other approaches will be vital to the development of practical systems.

3) *Spatial dimension and location dependence*: The radio-frequency spectrum is space–time–frequency varying. Thus, in a reasonable size cognitive radio network, the spectrum state will inevitably be different at different parts of the

network. How to acquire location information and fully exploit spatial diversity are crucial for interference management as well as maximal exploitation of spectrum opportunities.

In addition to these design issues, issues that require further attention are the coexistence of various heterogeneous SU networks and computational complexity and energy- and bandwidth efficiency of multiuser cooperative spectrum sensing and access policies. Furthermore, time synchronization and temporal allocation of resources in finding and accessing idle spectrum are important issues in managing the complicated problem of flexible spectrum use.

Finally, we note that, in this article, we could only provide short introductions to the many interesting and important techniques for spectrum sensing and access. More comprehensive treatments of the different techniques introduced in this article can be found in [3], [5], and [17].

AUTHORS

Jarmo Lundén (jarmo.lunden@aalto.fi) is with the Department of Signal Processing and Acoustics, School of Electrical Engineering, Aalto University, Finland.

Visa Koivunen (visa.koivunen@aalto.fi) is with the Department of Signal Processing and Acoustics, School of Electrical Engineering, Aalto University, Finland.

H. Vincent Poor (poor@princeton.edu) is with the Department of Electrical Engineering, Princeton University, New Jersey.

REFERENCES

[1] S. H. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of myopic sensing in multi-channel opportunistic access," *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 4040–4050, Sept. 2009.

[2] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multi-armed bandit problem," *Mach. Learn.*, vol. 47, no. 2–3, pp. 235–256, 2002.

[3] T. Alpcan, H. Boche, M. Honig and H.V. Poor, *Mechanisms and Games for Dynamic Spectrum Allocation*. Cambridge, UK: Cambridge Univ. Press, 2014.

[4] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio: State-of-the-art and recent advances," *IEEE Signal Processing Mag.*, vol. 29, no. 3, pp. 101–116, May 2012.

[5] E. Biglieri, A. Goldsmith, L. Greenstein, N. Mandayam, and H. V. Poor, *Principles of Cognitive Radio*. Cambridge, UK: Cambridge Univ. Press, 2013.

[6] B. E. Brodsky and B. S. Darkhovsky, "Minimax sequential tests for many composite hypothesis. I," *Theory Probab. Appl.*, vol. 52, no. 4, pp. 565–579, 2008.

[7] L. Bușoniu, R. Babuška, and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 38, no. 2, pp. 156–172, Mar. 2008.

[8] O. Cappé, A. Garivier, O.-A. Maillard, R. Munos, and G. Stoltz, "Kullback-Leibler upper confidence bounds for optimal sequential allocation," *Ann. Stat.*, vol. 41, no. 3, pp. 1516–1541, June 2013.

[9] S. Chaudhari, V. Koivunen, and H. V. Poor, "Autocorrelation-based decentralized sequential detection of OFDM signals in cognitive radios," *IEEE Trans. Signal Processing*, vol. 57, no. 7, pp. 2690–2700, July 2009.

[10] S. Chaudhari, J. Lundén, V. Koivunen, and H. V. Poor, "Cooperative sensing with imperfect reporting channels: Hard decisions or soft decisions?," *IEEE Trans. Signal Processing*, vol. 60, no. 1, pp. 18–28, Jan. 2012.

[11] Y. Chen, "Optimum number of secondary users in collaborative spectrum sensing considering resource usage efficiency," *IEEE Commun. Lett.*, vol. 12, no. 12, pp. 877–879, Dec. 2008.

[12] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 2053–2071, May 2008.

[13] J. C. Gittins and D. M. Jones, "A dynamic allocation index for the sequential design of experiments," in *Progress in Statistics*, J. Gani, Ed. Amsterdam: North Holland, 1974, pp. 241–266.

[14] O. Habachi, R. El-azouzi, and Y. Hayel, "A Stackelberg model for opportunistic sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2148–2159, May 2013.

[15] O. Hadjiliadis, H. Zhang, and H. V. Poor, "One shot schemes for decentralized quickest change detection," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3346–3359, July 2009.

[16] E. A. Hansen, D. S. Bernstein, and S. Zilberstein, "Dynamic programming for partially observable stochastic games," in *Proc. 19th Nat. Conf. Artificial Intelligence*, San Jose, CA, USA, July 25–29, 2004, pp. 709–715.

[17] E. Hossain, D. Niyato, and Z. Han, *Dynamic Spectrum Access and Management in Cognitive Radio Networks*. Cambridge, UK: Cambridge Univ. Press, 2009.

[18] A. M. Hussain, "Multisensor distributed sequential systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, no. 3, pp. 698–708, July 1994.

[19] H. Jiang, L. Lai, R. Fan, and H. V. Poor, "Optimal selection of channel sensing order in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 297–307, Jan. 2009.

[20] L. P. Kaebling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artif. Intell.*, vol. 110, no. 1–2, pp. 99–134, May 1998.

[21] S.-J. Kim and G. B. Giannakis, "Sequential and cooperative sensing for multi-channel cognitive radios," *IEEE Trans. Signal Processing*, vol. 58, no. 8, pp. 4239–4253, Aug. 2010.

[22] L. Lai, Y. Fan, and H. V. Poor, "Quickest detection in cognitive radio: A sequential change detection framework," in *Proc. IEEE Global Communications Conf. (GLOBECOM)*, New Orleans, LA, Nov. 30–Dec. 4, 2008, pp. 1–5.

[23] T. L. Lai, "Sequential analysis: Some classical problems and new challenges," *Statist. Sinica*, vol. 11, no. 2, pp. 303–408, Apr. 2001.

[24] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Adv. Appl. Math.*, vol. 6, no. 1, pp. 4–22, Mar. 1985.

[25] W.-Y. Lee and I. F. Akyildiz, "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3845–3857, Oct. 2008.

[26] K. B. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proc. IEEE*, vol. 97, no. 5, pp. 878–893, May 2009.

[27] H. Li, "Restless watchdog: Selective quickest spectrum sensing in multichannel cognitive radio systems," *EURASIP J. Adv. Signal Processing—Special Issue on Dynamic Spectrum Access for Wireless Networking*, vol. 2009, no. 9, pp. 1–12, Sept. 2009.

[28] H. Liu, K. Liu, and Q. Zhao, "Learning in a changing world: Restless multiarmed bandit with unknown dynamics," *IEEE Trans. Inform. Theory*, vol. 59, no. 3, pp. 1902–1916, Mar. 2013.

[29] K. Liu and Q. Zhao, "Indexability of restless bandit problems and optimality of Whittle index for dynamic multichannel access," *IEEE Trans. Inform. Theory*, vol. 56, no. 11, pp. 5547–5567, Nov. 2010.

[30] K. Liu, Q. Zhao, and B. Krishnamachari, "Dynamic multichannel access with imperfect channel state detection," *IEEE Trans. Signal Processing*, vol. 58, no. 5, pp. 2795–2808, May 2010.

[31] K. J. R. Liu and B. Wang, *Cognitive Radio Networking and Security*. Cambridge, UK: Cambridge Univ. Press, 2011.

[32] B. F. Lo and I. F. Akyildiz, "Reinforcement learning for cooperative sensing gain in cognitive radio ad hoc networks," *Wireless Netw.*, vol. 19, no. 6, pp. 1237–1250, Aug. 2013.

[33] J. Lundén, V. Koivunen, A. Huttunen, and H. V. Poor, "Collaborative cyclostationary spectrum sensing for cognitive radio systems," *IEEE Trans. Signal Processing*, vol. 57, no. 11, pp. 4182–4195, Nov. 2009.

[34] J. Lundén, S. R. Kulkarni, V. Koivunen, and H. V. Poor, "Multiagent reinforcement learning based spectrum sensing policies for cognitive radio networks," *J. Select. Topics Signal Processing*, vol. 7, no. 5, pp. 858–868, Oct. 2013.

[35] J. Ma, G. Zhao, and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4502–4507, Nov. 2008.

[36] S. Maleki, S. P. Chepuri, and G. Leus, "Optimization of hard fusion based spectrum sensing for energy-constrained cognitive radio networks," *Phys. Commun.*, vol. 9, pp. 193–198, Dec. 2013.

[37] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *Ann. Stat.*, vol. 14, no. 4, pp. 1379–1387, 1986.

[38] J. Nash, "Noncooperative games," *Ann. of Math.*, vol. 54, no. 2, pp. 286–295, Sept. 1951.

[39] J. Niño-Mora, "A $(2/3)n^3$ fast-pivoting algorithm for the Gittins index and optimal stopping of a Markov chain," *INFORMS J. Comput.*, vol. 19, no. 4, pp. 596–606, 2007.

[40] J. Oksanen, V. Koivunen, and H. V. Poor, "A sensing policy based on confidence bounds and a restless multi-armed bandit model," in *Proc. 46th Asilomar Conf. Signals, Systems and Computers (ASILOMAR)*, Pacific Grove, CA, Nov. 4–7, 2012, pp. 318–323.

[41] J. Oksanen, J. Lundén, and V. Koivunen, "Characterization of spatial diversity in cooperative spectrum sensing," in *Proc. 4th Int. Symp. Communications, Control, and Signal Processing (ISCCSP)*, Limassol, Cyprus, Mar. 3–5, 2010, pp. 1–5.

[42] J. Oksanen, J. Lundén, and V. Koivunen, "Reinforcement learning based sensing policy optimization for energy efficient cognitive radio networks," *Neurocomputing—Special Issue on Machine Learning for Signal Processing 2010*, vol. 80, pp. 102–110, Mar. 2012.

[43] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1–2, pp. 100–115, 1954.

[44] E. C. Y. Peh, Y.-C. Liang, Y. L. Guan, and Y. Zeng, "Optimization of cooperative sensing in cognitive radio networks: A sensing-throughput tradeoff view," *IEEE Trans. Veh. Technol.*, vol. 58, no. 9, pp. 5924–5299, Nov. 2009.

[45] M. Pollak, "A robust changepoint detection method," *Sequential Anal.*, vol. 29, no. 2, pp. 146–161, Apr. 2010.

[46] H. V. Poor and O. Hadjiliadis, *Quickest Detection*. Cambridge, UK: Cambridge Univ. Press, 2009.

[47] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: A low-communication-rate scheme for distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[48] G. A. Rummery and M. Niranjan, "On-line Q-learning using connectionist systems," Tech. Rep. CUED/F-INFENG/TR 166, Cambridge Univ., Dept. Engineering, England, Sept. 1994.

[49] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Başar, "Coalitional game theory for communication networks: A tutorial," *IEEE Signal Processing Mag.*, vol. 26, no. 5, pp. 77–97, Sept. 2009.

[50] W. Saad, Z. Han, R. Zheng, A. Hjørungnes, T. Başar, and H. V. Poor, "Coalitional games in partition form for joint spectrum sensing and access in cognitive radio networks," *IEEE J. Select. Topics Signal Processing*, vol. 6, no. 2, pp. 195–209, Apr. 2012.

[51] S. Sardellitti, M. Giona, and S. Barbarossa, "Fast distributed average consensus algorithms based on advection-diffusion processes," *IEEE Trans. Signal Processing*, vol. 58, no. 2, pp. 826–842, Feb. 2010.

[52] G. Scutari and J.-S. Pang, "Joint sensing and power allocation in nonconvex cognitive radio games: Nash equilibria and distributed algorithms," *IEEE Trans. Inform. Theory*, vol. 59, no. 7, pp. 4626–4661, July 2013.

[53] A. N. Shiryaev, "On optimum methods in quickest detection problems," *Theory Probab. Appl.*, vol. 8, no. 1, pp. 22–46, Jan. 1963.

[54] S. P. Singh, T. Jaakkola, M. L. Littman, and C. Szepesvári, "Convergence results for single-step on-policy reinforcement-learning algorithms," *Mach. Learn.*, vol. 39, no. 3, pp. 287–308, Mar. 2000.

[55] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable Markov processes over a finite horizon," *Oper. Res.*, vol. 21, no. 5, pp. 1071–1088, Sept.–Oct. 1973.

[56] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 1998.

[57] S. Tantaratana and H. V. Poor, "Asymptotic efficiencies of truncated sequential tests," *IEEE Trans. Inform. Theory*, vol. 28, no. 6, pp. 911–923, Nov. 1982.

[58] J. Unnikrishnan, V. V. Veeravalli, and S. Meyn, "Minimax robust quickest change detection," *IEEE Trans. Inform. Theory*, vol. 57, no. 3, pp. 1604–1614, Mar. 2011.

[59] V. V. Veeravalli, "Decentralized quickest change detection," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1657–1665, May 2001.

[60] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part I—Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[61] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Ann. Math. Stat.*, vol. 19, no. 3, pp. 326–339, 1948.

[62] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, nos. 3–4, pp. 279–292, May 1992.

[63] P. Whittle, "Restless bandits: Activity allocation in a changing world," *J. Appl. Prob.*, vol. 25, pp. 287–298, 1988.

[64] F. Zeng, C. Li, and Z. Tian, "Distributed compressive spectrum sensing in cooperative multihop cognitive networks," *IEEE Select. Topics Signal Processing*, vol. 5. no. 1, pp. 37–48, Feb. 2011.

[65] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE J. Select. Areas Commun.*, vol. 25, no. 3, pp. 589–600, Apr. 2007.

[66] Q. Zhao, B. Krishnamachari, and K. Liu, "On myopic sensing for multi-channel opportunistic access: Structure, optimality, and performance," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5431–5440, Dec. 2008.

[67] Q. Zhao and J. Ye, "Quickest detection in multiple on-off processes," *IEEE Trans. Signal Processing*, vol. 58, no. 12, pp. 5994–6006, Dec. 2010.

[68] Q. Zou, S. Zheng, and A. H. Sayed, "Cooperative sensing via sequential detection," *IEEE Trans. Signal Processing*, vol. 58, no. 12, pp. 6266–6283, Dec. 2010.

[SP]

[Nicholas D. Sidiropoulos and Efthymios E. Tsakonas]

Signal Processing and Optimization Tools for Conference Review and Session Assignment

[How to ease the burden on technical chairs and improve review quality at the same time]

Anyone who has served as a technical program committee (TPC) chair for a conference (or program manager for a funding agency) understands that paper (or proposal panel) review assignment is a demanding job that takes a lot of time, and reviewers are rarely satisfied with the end results. This article presents signal processing tools for two critical "mass assignment" tasks: assigning papers (or proposals) to reviewers in a way that matches reviewing expertise to scientific content while respecting the reviewers' capacity constraints and splitting accepted papers (or submitted proposals) to sessions (panels) while adhering to session (panel) capacity constraints. The basic idea is to use feature vectors to represent papers and reviewers. Features can be key words or phrases (e.g., *optimization* or *sensor networks*) or other types of attributes (e.g., *timeliness*). This viewpoint enables optimal assignment problem formulations that make sense from a scientific and practical point of view. While optimal solutions are hard to compute for a large number of papers and

IMAGE LICENSED BY GRAPHIC STOCK

reviewers, high-quality approximate solutions of moderate complexity are developed here using familiar signal processing and optimization tools. These algorithmic solutions easily outperform days of expert manual work as demonstrated in experiments with real conference data.

The credibility of our scientific enterprise relies heavily on the peer-review system. Whereas many contributions are eventually still individually judged (e.g., when submitted for journal publication), there are at least two important modes of mass peer review at the center stage of scientific innovation: proposal review panels and conference reviewing. Paper and proposal review assignment is a difficult and tedious job that takes a lot of time, and, despite good intentions, often results in some awkward assignments.

A TPC chair's job includes 1) assigning reviewers to each paper, making every effort to match reviewing expertise to paper content while respecting the reviewer capacity constraints; 2) reading the submitted reviews and making an accept/reject decision for each paper, keeping in mind the target acceptance rate and the number of papers that can be presented at the conference; and 3) splitting the accepted papers into sessions, such that each session has a

[FIG1] An example of what usually happens when one tries to maximize reviewer satisfaction (or similarity score) alone: Both green and red assignments have the same affinity score, but only the assignment in red ensures a scientifically sound paper review.

coherent theme, while adhering to session capacity constraints. The latter is the paper-to-session assignment problem. A program manager's job likewise includes 1) splitting the list of submitted proposals into smaller thematic batches to be assigned to review panels while adhering to panel capacity constraints (the proposal-to-panel assignment problem), 2) selecting reviewers to invite for each panel, and 3) assigning panelists to each proposal, trying to match reviewing expertise to the proposal content while respecting the panelist capacity constraints.

Given the difficulty and effort it takes to effectively solve these assignment problems, it is hard to believe that most TPC chairs and many program managers still operate without using the appropriate algorithmic aids to get the job done faster and better. There are two main reasons for this: 1) it is hard for a machine to nail down the essence of a submitted research paper or proposal and make a scientifically sound call on what is an appropriate set of reviewers, and 2) conference- or program-specific constraints require custom coding.

Generic computerized assignment algorithms (e.g., Cyberchair) are available, but these rely on reviewing "bids" or preference ratings, or a scalar similarity score between the contents of each paper and the expertise of each reviewer. Given the similarity (or affinity) paper-reviewer matrix, an assignment that maximizes the total affinity can be formulated as an integer linear programming problem. This formulation has been shown to be a totally unimodular program, which implies that an optimal solution can be computed at a modest complexity [1]; see also [2]. Using reviewer preferences for assignment certainly keeps the reviewers happy; however, it has two important pitfalls.

1) Each paper or proposal usually requires multiple types of expertise for proper review. For example, a paper on cross-layer resource allocation in wireless networking requires expertise in physical layer wireless communication, optimization, and networking. Using an aggregate preference or similarity score per reviewer can (and does) result in assignments where no reviewer covers a certain aspect of the paper (e.g., networking). This is, of course, highly undesirable, as already noted in some earlier work on automated review assignment [1], [3], [4]. A typical situation is depicted in Figure 1, which clearly shows the deficiency of total similarity/affinity score-based assignments. [While it is conceptually possible that one might be able to judiciously design a paper-reviewer score matrix that prohibits such bad assignments when used in conjunction with the totally unimodular programming approach in [1] and [2], this seems like a daunting task. Entry (p, r) of such a matrix should not only depend on the feature vectors of paper p and reviewer r; it should be a function of the feature vectors of potentially all papers and all reviewers.]

2) Reviewers tend to down-weight past experience in favor of their current interests when clicking on topical areas to summarize their expertise and generally bid to review papers or proposals that are close to their current interests, "in fashion," or from well-known researchers, without regard to the collective reviewing needs of the conference or panel. The TPC chair or program manager often has to tap a reviewer's past expertise to
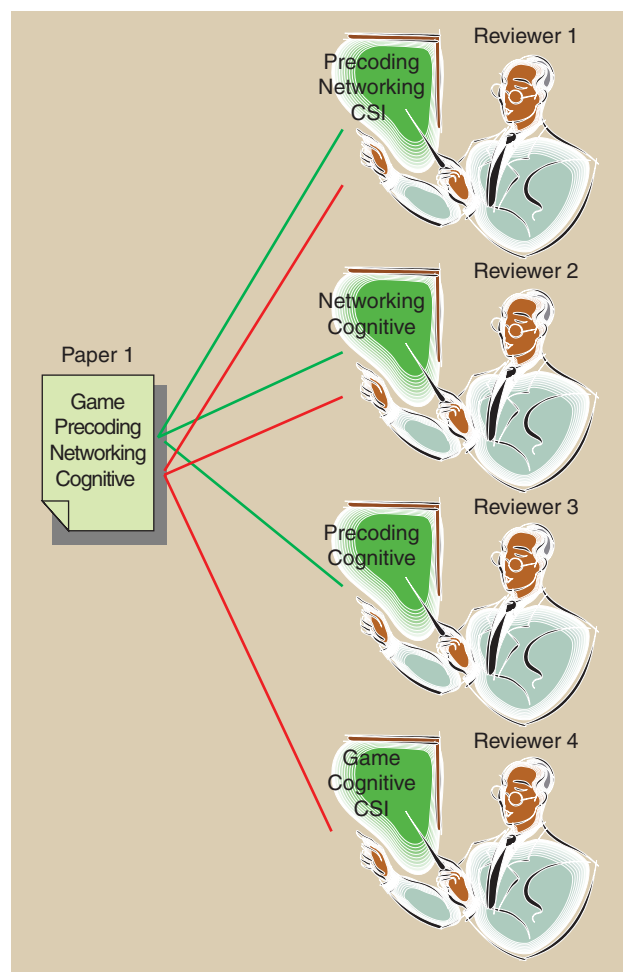
VISUALIZING PAPERS AND REVIEWERS

The number of key words/features used to describe papers and reviewers will typically be in the order of dozens, making it hard to visualize the distribution of papers and reviewers in a feature space. One approach is to compute the first two or three principal components and project those points onto the principal subspace for visualization. Another tool that is commonly used for visualization is multidimensional scaling (MDS). Given a matrix of pairwise distances between m objects, MDS computes a map of m points in two-dimensional (2-D) [or three-dimensional (3-D)] space that approximately preserves the given distances.

Figure S1 shows 2-D MDS maps of points corresponding to papers and reviewers from the Signal Processing for Communications and Networking Technical Committee (SPCOM TC) track of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2009, for which Nicholas D. Sidiropoulos served as TPC chair. The dimension of the vector space is $N = 44$—that is, there are 44 key words, and each paper or reviewer is represented by a (sparse) 44×1 vector. Figure S1 shows a map of (a) papers, (b) reviewers, and (c) a joint map of papers and reviewers. Notice how papers are clustered in (a), but this is not evident from the joint map (c). The reviewers are almost uniformly scattered, which speaks for the difficulty of optimal assignment: real data do not nicely fall in clusters. This situation is typical in our experience.
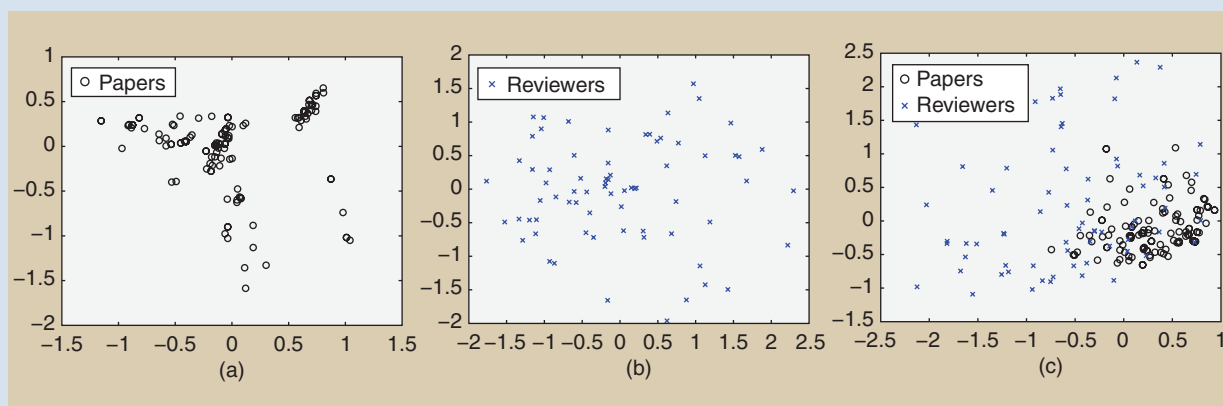
[FIGS1] (a) An MDS visualization of papers, (b) reviewers, and (c) joint papers and reviewers.

ensure a fair and unbiased assignment to the extent possible. These factors are very difficult to capture by reviewing preferences or aggregate similarity scores.

The first step toward a more pragmatic approach is a multidimensional description of each reviewer and each paper or proposal, in a common feature space that captures the essential dimensions of expertise for the specific scientific domain. In other words, we advocate viewing reviewers and papers/proposals as points in a higher-dimensional vector space. The canonical coordinates in this vector space are key words or phrases used to represent papers and reviewers (e.g., *optimization* or *sensor networks*), or other types of attributes (e.g., *timeliness*). This concept is illustrated in Figures 2 and 3, and it is central to our approach (see also "Visualizing Papers and Reviewers"). Note that feature vectors are widely used in the machine-learning literature; see, e.g., [5] and [6].

The list of keys for the papers (dimensions of the feature vector) can be produced as follows:

■ The list can be prepared by the TPC chair before submission, in which case authors can mark the features relevant to their paper at the time of submission. This would correspond to a refined Editors' Information Classification Scheme.

■ They can be compiled by taking the union of standard plus free-text key words provided by the authors at submission time, followed by stemming to consolidate synonyms.

■ They can be parsed from the list of submitted paper titles.

This parsing can be done manually by the TPC chair (for up to a few hundred papers—a seasoned chair can process about three papers per minute), or it can be automated using text retrieval [7] and consolidation tools [8]. Natural language processing will likely be helpful in this context, but this remains to be seen in practice. At any rate, spending a couple of hours producing a list of keys and marking papers is far less than what is needed for producing a well-rounded technical program from the list of accepted papers, let alone producing a scientifically sound review assignment.

■ Most conferences and workshops recur annually or periodically; therefore, a prepared list of key words for the previous edition can serve as an excellent starting point for the next one, with the addition of a few key words for emerging topics and possible deletion or consolidation of those that are obsolete.

Drawing upon this multidimensional description of papers and reviewers, this article aims to present signal processing tools for paper-to-reviewer assignment and paper-to-session assignment. We examine these two problems in the remainder of this article.

A PRIOR ART

PAPER-TO-REVIEWER ASSIGNMENT

In addition to the key works [1], [9] and related follow-up work, such as [2], there are several more references on mass review

assignment, e.g., [10] and the references therein. Those that are related to our viewpoint are reviewed in this article. Our vector space viewpoint of review assignment is implicit in [3], which considered representing each reviewer and each proposal with a list of key words or terms in a common term space and proposed evaluating reviewing assignments and making additional reviewer recommendations by measuring how the assigned reviewers collectively cover a proposal's key words; see also later work in [4]. The work of Hettich and Pazzani [3] is in fact a lucid and very insightful account of lessons learned in designing and implementing an early review aide system at the National Science Foundation (NSF) several years ago. What is missing from [3] (and [4]) is formulating review assignment as a joint optimization problem subject to reviewing capacity constraints, addressing complexity issues, and coming up with suitable algorithms to solve it. Instead, a simple greedy hill-climbing approach to making individual recommendations one reviewer at a time is discussed in [3]. Paper-to-session/proposal-to-panel assignment is not discussed at all in [3] and [4].

Today, several NSF program managers use a tool developed in [11] for review assignment. The approach in [11] is based on panelist reviewing preferences and uses a generalized assignment formulation with a branch-and-bound solution technique that is complex for large problems; however, it is tailored for the NSF panel review and complexity is not a major issue for modest panel sizes. On the other hand, it does not account for the need to cover all bases in reviewing a particular proposal or the bias that is typical in reviewing preferences. Additional work related to review assignment can be found in [12]; see also [13] for a recent overview of assignment problems.

PAPER-TO-SESSION ASSIGNMENT

Fitting the accepted papers into sessions is a clustering problem under equality constraints on the number of points per cluster—because each session has a fixed capacity. In this article, we focus on clustering using a centroid model, in which each cluster is represented by a single mean vector, and we have a given number of data points per cluster. In our context, each cluster corresponds to a session, and its centroid reflects the key words that are dominant in that session, thereby serving as a crude session title (which can be polished later by the TPC chair). The traditional signal processing and computer science literature treats clustering mostly using the well-known k-means algorithm [14], which cannot be directly applied in our context due to the presence of the session capacity constraints. Modifications of k-means to account for must-link/cannot-link constraints are discussed in [15], distance-type constraints on the cluster centers are discussed in [16], and lower-bound constraints on the number of points per cluster are discussed in [17]. As an alternative to alternating optimization-based k-means, approximation algorithms based on convex (semidefinite) optimization [18] are also known; see, e.g., [19] and the references therein.

Our formulation of paper-to-session assignment can be called a *capacitated k-means* problem. Whereas the general literature on clustering is immense [20], [21], we did not find any prior work on capacitated k-means, likely because there is no motivation to specify cluster sizes a priori in most applications of unsupervised clustering—where we typically know little about the clusters we

are trying to find. Imposing a lower bound on cluster size may seem reasonable to avoid degeneracy, but an upper bound does not make sense in most other applications.

In practice, paper-to-reviewer assignment naturally precedes paper-to-session assignment. Paper-to-reviewer assignment is more challenging than paper-to-session assignment because there are typically many more papers submitted than accepted and many more reviewers than sessions in the final program. Furthermore, paper-to-reviewer assignment quality is more important from a scientific and ethical point of view. Yet the paper-to-session assignment problem is important and hard in its own right (we will show that it is NP-hard, in fact). There is also something special about the paper-to-session assignment problem: it is near and dear to our signal processing hearts. We will show how to modify k-means to account for strict cluster capacity constraints and produce a very practical and efficient low-complexity algorithm. We will also develop a more sophisticated one-shot approximation that can be used in smaller paper-to-session assignment problem instances. For these reasons, and despite the conceptual order of the two problems, we will start from the paper-to-session assignment problem. Before proceeding to the mathematical formulations, we first briefly review the mathematical tools that will be used.

MATHEMATICAL PRELIMINARIES

Assignment problems are optimization problems of a combinatorial nature; some have a special structure that enables efficient solution, while others are provably hard, even though they may not look all that different at first sight. The good news is that some of these problems can be well approximated (albeit not optimally solved) using convex optimization tools.

One way to deal with an optimization problem that is hard to solve is to efficiently obtain an approximate solution through convex relaxation. This comprises two steps (if the cost function of the original problem is not convex, then an additional transformation is required). In the first step, one replaces the feasible region of the original problem with a convex superset (hence the term *relaxation*); then the resulting problem is solved using convex optimization algorithms. In the second step, one converts the solution of the relaxed problem into a good admissible solution for the original problem through suitable postprocessing. The postprocessing step involves projection of the solution of the relaxed problem (and possibly related candidates generated via randomization) onto the feasible set of the original problem. Obviously, the optimal value of the relaxed problem provides a bound on the optimal value of the original problem; one goal is to find the tightest such bound (make the relaxation as tight as possible), as this impacts the quality of the final solution. We now illustrate how the idea of convex relaxation applies to both paper-to-session and review assignment.

RELAXATION OF PAPER-TO-SESSION ASSIGNMENT

The main algorithm is given in the section "Proposed Algorithm for Paper-to-Session Assignment" and is based on alternating optimization; see [16] and the references therein. This is an iterative procedure for optimizing a cost function by alternating

conditional updates of different subsets of variables given the rest of the variables. However, we also show in the section "Gauging the Optimality Gap: Semidefinite Relaxation" that paper-to-session assignment can be equivalently rewritten as a quadratically constrained quadratic program (QCQP). This has the form

$$\underset{x}{\text{minimize}} \ \mathbf{x}^\mathrm{T}\mathbf{Q}\mathbf{x}$$
$$\text{subject to: } \mathbf{x}^\mathrm{T}\mathbf{C}_i\mathbf{x} \le b_i, \ i = 1,\dots,n+1, \qquad (1)$$

with $\mathbf{Q}$ and $\mathbf{C}_i \in \mathbb{R}^{n\times n}$ symmetric matrices and b_i scalar quantities. Casting paper-to-session assignment as a QCQP is interesting since there are many tools available in the literature for quadratic optimization and they are well understood. The best convex relaxation bounds for (1) are based on semidefinite relaxation (SDR) [22]: one starts by 1) rewriting the quadratic cost in (1) as $\mathrm{Tr}(\mathbf{x}^\mathrm{T}\mathbf{Q}\mathbf{x}) = \mathrm{Tr}(\mathbf{Q}\mathbf{x}\mathbf{x}^\mathrm{T})$ (and similarly rewriting every quadratic constraint), and then 2) lifting the problem in a higher dimensional space using the change of variables $\mathbf{X} = \mathbf{x}\mathbf{x}^\mathrm{T}$. This lifting isolates the nonconvexities of the original QCQP into a single rank-1 constraint. The rank-1 constraint is subsequently relaxed into a convex, positive semidefinite cone constraint [18], or even simply dropped, thereby producing a convex (relaxed) problem. This is the main idea of SDR—the details of the transformation along with the corresponding postprocessing step, which produces the final approximate solution, are described in the section "Gauging the Optimality Gap: Semidefinite Relaxation."

PAPER-TO-REVIEWER ASSIGNMENT

As we explain in detail in the section "The Review and Assignment Problem," the associated optimization problem has the following form:

$$\underset{x}{\text{minimize}} \ f(\mathbf{x})$$
$$\text{subject to: } \mathbf{A}\mathbf{x} \preceq \mathbf{b}, \ \mathbf{x} \in \{0,1\}^n, \qquad (2)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a convex piecewise linear function in the variables $\mathbf{x} \in \mathbb{R}^n$, and $\preceq$ indicates componentwise inequality. The set defined by the inequality $\mathbf{A}\mathbf{x} \preceq \mathbf{b}$ is convex and is called a *polyhedron* [18]. Note that, even though the cost in (2) is convex, the design variables are Boolean, either zero or one. Boolean constraints are nonconvex constraints; in fact, it is often convenient to write them explicitly as quadratic equalities since $\mathbf{x}_i \in \{0,1\} \iff \mathbf{x}_i(1 - \mathbf{x}_i) = 0$.

In the first step, we produce the tightest convex relaxation (to be concise, the phrase "tightest convex relaxation" should be interpreted as "tightest relaxation in the class of Langragian relaxations"; see [18]) of (2): it can be shown that this is tantamount to replacing the Boolean constraints on the $\mathbf{x}_i$s with the interval ones $0 \le \mathbf{x}_i \le 1$ [18, Ch. 5]. We refer to this relaxation as *linear programming relaxation* because the resulting problem can be cast as a linear program (LP). Since the relaxed solution is not guaranteed to be Boolean, in the second step (the postprocessing), we make use of the structure of $\mathbf{A}$ and the nature of $\mathbf{b}$ to efficiently compute the Euclidean projection of the relaxed solution onto the feasible set of (2). This is the main idea—we defer the details to the section "The Review Assignment Problem."

PAPER-TO-SESSION ASSIGNMENT: A CLOSER LOOK

Recall from Figure 2 that we use feature vectors to represent papers. If N is the number of features (key words), then feature vectors are nonnegative vectors in $\mathbb{R}^N$. Let $\mathbf{p}_i$ be the $N \times 1$ feature vector for paper $i \in \mathcal{I} \triangleq \{1,\dots,I\}$, where I is the total number of
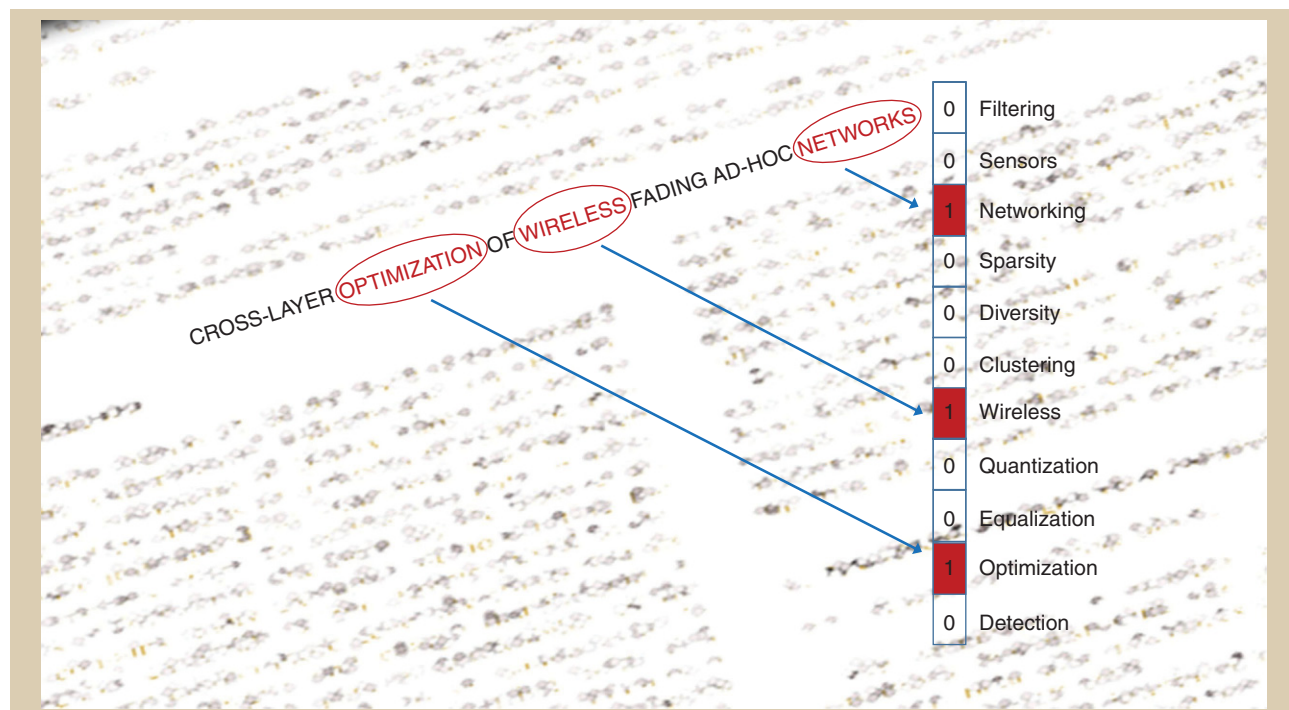
[FIG2] Representing a paper as a point (feature vector) in the key word space. In this illustration, the feature vector is Boolean, with 1 if the paper possesses the specific key word and 0 otherwise.

(accepted) papers. Define the $N \times I$ matrix $\mathbf{P} := [\mathbf{p}_1, \cdots, \mathbf{p}_I]$. The capacity of session $j \in \mathcal{J} \triangleq \{1, \cdots, J\}$ is denoted c_j; $\sum_{j=1}^{J} c_j = I$, i.e., the total number of accepted papers.

The design variables are the $N \times J$ matrix of session centers $\mathbf{S} := [\mathbf{s}_1, \cdots, \mathbf{s}_J]$, where $\mathbf{s}_j$ is the center (profile, or title) of session j; and the $J \times I$ paper-to-session assignment matrix $\mathbf{X}$. The elements $\mathbf{X}_{ji}$ of $\mathbf{X}$ must satisfy the following constraints:

$$\mathbf{X}_{ji} \in \{0, 1\}, \forall i \in \mathcal{I}, j \in \mathcal{J}, \tag{3a}$$

$$\sum_{j=1}^{J} \mathbf{X}_{ji} = 1, \forall i \in \mathcal{I}, \tag{3b}$$

$$\sum_{i=1}^{I} \mathbf{X}_{ji} = c_j, \forall j \in \mathcal{J}. \tag{3c}$$

Here, $\mathbf{X}_{ji} = 1$ means that paper i is assigned to session j. The constraint $\sum_{j=1}^{J} \mathbf{X}_{ji} = 1$ ensures that paper i will be assigned to the one and only one session, whereas $\sum_{i=1}^{I} \mathbf{X}_{ji} = c_j$ enforces the capacity constraint for session j.

For brevity, let $\mathcal{A}$ denote the set of matrices $\mathbf{X} \in \mathbb{R}^{J \times I}$ that satisfy (3a)–(3c). With these definitions, the paper-to-session (or technical program optimization) problem can be posed as follows: assign papers to sessions (pick $\mathbf{X}$) and find the appropriate "session titles" (pick $\mathbf{S}$) to

$$\underset{\mathbf{S}, \mathbf{X}}{\text{minimize}} \, \| \mathbf{P} - \mathbf{SX} \|_F^2, \tag{4a}$$

$$\text{subject to: } \mathbf{X} \in \mathcal{A}. \tag{4b}$$

See "Distance Considerations" for a discussion on the choice of distance measure.

A property worth pointing out explicitly is that any matrix feasible for (4) is row-orthogonal. To see this, define the vector of session capacities $\mathbf{c} = [c_1, c_2, \cdots, c_J]^T$ and the $J \times J$ matrix $\mathbf{\Lambda} = \text{Diag}(\mathbf{c})$, with the entries of the vector $\mathbf{c}$ on its main diagonal and zero elsewhere. Then, we have that

$$\mathbf{X} \in \mathcal{A} \implies \mathbf{X}\mathbf{X}^T = \mathbf{\Lambda}. \tag{5}$$

This observation will be useful on multiple occasions later on, in the problem transformations.

REMARK 1
Note that, in principle, one can place inequality constraints on the session capacities instead of the equality constraints (3c). Inequality constraints on the capacities make sense perhaps for poster sessions, but not for oral sessions, where a fixed number of papers should be presented. Although using inequalities instead of equalities is possible, the overall treatment of the problem (in particular, the material in the section "Gauging the Optimality Gap: Semidefinite Relaxation") becomes more involved. We choose to work with equality constraints to simplify exposition; after all, the TPC chair can explore minor reallocations of poster session capacities by running the proposed algorithms a few times if so desired. Also note that collisions (an author having to present simultaneously in two parallel sessions) are usually handled at the end by permuting the order of the presentation of papers in oral sessions or manual reallocation to a different session if a poster presentation is involved. Such scheduling conflicts are usually rare and also depend on the metadata, such as who is the presenting coauthor, and session time-scheduling, which in turn depends on the session content, the number of parallel tracks, room

---

DISTANCE CONSIDERATIONS
Returning to (4),

$$\min_{\mathbf{S}, \mathbf{X} \in \mathcal{A}} \| \mathbf{P} - \mathbf{SX} \|_F^2 \iff \min_{\mathbf{X} \in \mathcal{A}} \sum_{j=1}^{J} \min_{\mathbf{s}_j} \sum_{i \in I_j(\mathbf{X})} \| \mathbf{p}_i - \mathbf{s}_j \|_2^2,$$

the use of the Euclidean distance can be motivated as follows. Assume that the $\mathbf{p}_i$'s are drawn from J classes, with each class represented by a class mean, $\mathbf{s}_j$. A paper drawn from class j follows a multivariate Gaussian distribution $\mathcal{N}(\mathbf{s}_j, \sigma^2 \mathbf{I})$. Different papers are independently distributed, and we know the number of papers in each class (the session capacities). Then, maximum likelihood joint paper classification and class mean estimation reduces to the above formulation, as can be easily seen by taking the log-likelihood and invoking independence.

The Gaussian assumption/Euclidean distance can be motivated in many ways; a testament to its ubiquity is that classical k-means uses Euclidean distance. But there are many alternatives that might be worth investigating. If clusters appear to be oriented, then a Mahalanobis distance (quadratic form involving the inverse cluster covariance matrix) is more appropriate, but the cluster covariance(s) should be estimated as well. If the $\mathbf{p}_i$'s can be modeled as probability mass functions, then the Kullback–Leibler divergence can be well-motivated; see also [40] for a tutorial overview of clustering with Bregman divergences.

In our numerical experiments, we have limited ourselves to using binary feature vectors, mainly because this is enough to capture the essence of the problems considered. Richer alphabets are needed to capture the degree of expertise required in each latent dimension—some papers may only need common expertise in a particular area, while others may demand much deeper understanding. If we stay with binary features, however, then a more natural metric is the Hamming distance $d_H(\mathbf{p}_i, \mathbf{s}_j) = \sum_{n=1}^{N} 1(\mathbf{p}_i(n) \neq \mathbf{s}_j(n))$. This corresponds to saying that the probability of drawing $\mathbf{p}_i$ from class j is $q^{d_H(\mathbf{p}_i, \mathbf{s}_j)}(1-q)^{N-d_H(\mathbf{p}_i, \mathbf{s}_j)}$, for some $q < 0.5$, so the more likely vectors are those with few bit flips. If we also force the estimated $\mathbf{s}_j$s to be 0-1 binary, the Hamming distance reduces to ℓ_1-distance, i.e., the sum of absolute values. Then, the conditional update of each $\mathbf{s}_j$ is the elementwise median of the vectors in the cluster. Although not shown here, in many of these variations, the conditional update of $\mathbf{X}$ given $\mathbf{S}$ is also tractable, i.e., it reduces to a totally unimodular LP.

The appropriateness of any assumption and engineering design is ultimately judged by how well it performs in practice. Euclidean distance works well enough in our context, as illustrated in our experiments with real conference data.

capacities, etc. While it is possible to incorporate some of these aspects in the problem formulation, we prefer to keep the exposition simple and address the core problem instead.

COMPLEXITY OF OPTIMAL PAPER-TO-SESSION ASSIGNMENT

If we drop the session capacity constraints (3c) from (2), a classic k-means problem emerges. k-means is NP-hard; in loose terms, this means that we cannot expect to solve an arbitrary instance of k-means in time polynomial in the number of papers I. In the signal processing community, k-means is also known as *vector quantization* (VQ), usually dealt with using the celebrated (generalized) Lloyd–Max (GLM) [23], [24] or Linde–Buzo–Gray (LBG) algorithm [25], which is an alternating optimization procedure. The reason we usually resort to LBG is precisely because the problem is hard, and the LBG iteration offers an attractive simplicity-performance-complexity tradeoff. Proof that k-means is NP-hard was only recently provided [26], [27].

Here, we are actually dealing with a restriction of the VQ/k-means problem due to the session capacity constraints, which will always be active. We show next that, unfortunately, this restriction is also an NP-hard problem. Given a feasible $\mathbf{X}$, let $I_j(\mathbf{X})$ denote the indices of papers falling in session j. Then,

$$\min_{\mathbf{S},\mathbf{X}} \|\mathbf{P} - \mathbf{SX}\|_F^2 \iff \min_{\mathbf{X}}\left\{ \min_{\mathbf{S}} \|\mathbf{P} - \mathbf{SX}\|_F^2 \right\} \iff$$

$$\min_{\mathbf{X}} \sum_{j=1}^{J} \min_{\mathbf{s}_j} \sum_{i \in I_j(\mathbf{X})} \|\mathbf{p}_i - \mathbf{s}_j\|_2^2.$$

The solution of the inner minimization for $\mathbf{s}_j$ is clearly the mean of those vectors falling in session j. Setting $\mathbf{s}_j$ equal to this mean, i.e., setting $\mathbf{s}_j$ equal to

$$\mathbf{s}_j^* \triangleq \frac{1}{|I_j(\mathbf{X})|} \sum_{\ell \in I_j(\mathbf{X})} \mathbf{p}_\ell,$$

it can be easily shown by expanding the squares that

$$\sum_{i \in I_j(\mathbf{X})} \|\mathbf{p}_i - \mathbf{s}_j^*\|_2^2 = \frac{1}{2\,|I_j(\mathbf{X})|} \sum_{i \in I_j(\mathbf{X})} \sum_{k \in I_j(\mathbf{X})} \|\mathbf{p}_i - \mathbf{p}_k\|_2^2,$$

where $|\cdot|$ denotes cardinality. If all session capacities are equal, we may thus use the following criterion instead:

$$\min_{\mathbf{X}} \sum_{j=1}^{J} \sum_{i \in I_j(\mathbf{X})} \sum_{k \in I_j(\mathbf{X})} \|\mathbf{p}_i - \mathbf{p}_k\|_2^2,$$

which is to be optimized over $\mathbf{X} \in \mathcal{A}$. This is now what is known as the *minimum k-clustering sum* problem (in our context J plays the role of k), which is in the list of NP-hard problems [28]; see also [29]—The poor TPC chair souls were right all along.

Claim 1

Technical program optimization (paper-to-session assignment, capacitated k-means) is NP-hard.
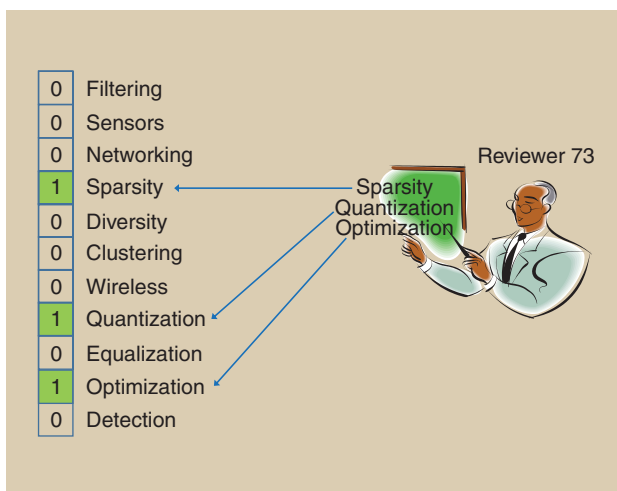
[FIG3] Representing a reviewer as a point (feature vector) in the same key word space. The feature vector in this particular illustration is Boolean.

The implication is that we cannot expect to solve an arbitrary instance of (2) in complexity polynomial in the number of papers I. It has been shown in [29] (see also [28]) that the minimum k-clustering sum problem can be approximated within a factor of 2—but the algorithm that provides this approximation guarantee has exponential complexity in J. Since J is not small in our context, we will instead explore familiar signal processing tools to obtain conceptually simple and performance-wise satisfactory solutions.

PROPOSED ALGORITHM FOR PAPER-TO-SESSION ASSIGNMENT

The GLM/LBG algorithm is typically used for VQ design. GLM/LBG alternates between optimizing the codebook $\mathbf{S}$ for a given assignment $\mathbf{X}$ and optimizing the assignment $\mathbf{X}$ for a given codebook $\mathbf{S}$. GLM/LBG exploits necessary optimality conditions, implying that $\mathbf{s}_j$ should be the mean of those $\mathbf{p}_i$s assigned to session j, and $\mathbf{p}_i$ should be assigned to the closest $\mathbf{s}_j$; these yield simple conditional updates. The GLM/LBG iteration converges in terms of fit, but the quality of the final solution depends heavily on the initialization.

GLM/LBG cannot be directly applied in our present context because of the presence of the session capacity constraints. In the following, we propose one possible iteration that explicitly takes these constraints into account.

Given a feasible assignment $\mathbf{X}$, the update for $\mathbf{S}$ is simple and, in fact, identical to the corresponding update in GLM/LBG. The step that requires closer scrutiny is the update of $\mathbf{X}$ given $\mathbf{S}$

$$\underset{\mathbf{X}}{\text{minimize}} \ \|\mathbf{P} - \mathbf{SX}\|_F^2 \tag{6a}$$

$$\text{subject to: } \mathbf{X} \in \mathcal{A}. \tag{6b}$$

Fortunately, it turns out that an optimal point for (6) can be computed easily, without having to search over all feasible assignments $\mathbf{X}$. To explain how this is possible, note first that the objective function in (6a) can be expressed as $\|\mathbf{P} - \mathbf{SX}\|_F^2 = \|\mathbf{P}\|_F^2 - 2\mathrm{Tr}(\mathbf{P}^T\mathbf{SX}) + \|\mathbf{SX}\|_F^2$, and observe that the quadratic term $\|\mathbf{SX}\|_F^2$

remains constant for any feasible assignment $\mathbf{X}$. This is because of the property in (5), since $\|\mathbf{SX}\|_F^2 = \mathrm{Tr}(\mathbf{X}^T\mathbf{S}^T\mathbf{SX}) = \mathrm{Tr}(\mathbf{XX}^T\mathbf{S}^T\mathbf{S}) = \mathrm{Tr}(\mathbf{\Lambda}\mathbf{S}^T\mathbf{S})$. Thus, the conditional update of $\mathbf{X}$ given $\mathbf{S}$ can be done by solving the Boolean LP

$$\underset{\mathbf{X}}{\text{maximize}} \quad \mathrm{Tr}(\mathbf{P}^T\mathbf{SX}) \qquad (7a)$$

$$\text{subject to: } \mathbf{X} \in \mathcal{A}. \qquad (7b)$$

Problem (7) is the so-called semiassignment problem, and there are many efficient algorithms for its solution. For example, the shortest augmenting path algorithm from [30] is applicable, which computes the solution of (7) at complexity $O(JI^2)$.

Although the shortest augmenting path algorithm from [30] is arguably one of the best choices (among the applicable algorithms) for carrying out the $\mathbf{X}$-update, we here also discuss how this can be done using linear programming. We believe that this discussion offers more insights and demonstrates an interesting connection between convex and combinatorial optimization. Observe first that the system of equations in (3b)–(3c) is linear and, therefore, can be written in the form $\mathbf{Gx} = \mathbf{d}$, where $\mathbf{x} \triangleq \mathrm{vec}(\mathbf{X})$. [The operation $\mathrm{vec}(\mathbf{X})$ stacks the columns of the matrix $\mathbf{X}$ into a vector.] Now, the coefficient matrix $\mathbf{G}$ is totally unimodular, i.e., every square submatrix has a determinant of value $0, \pm 1$; and $\mathbf{d}$ is a vector of integers. As a result [31], the polyhedron

$$0 \le \mathbf{X}_{ji} \le 1, \forall i \in \mathcal{I}, j \in \mathcal{J}$$

$$\sum_{j=1}^{J} \mathbf{X}_{ji} = 1, \forall i \in \mathcal{I}$$

$$\sum_{i=1}^{I} \mathbf{X}_{ji} = c_j, \forall j \in \mathcal{J}$$

is the convex hull of all assignments $\mathbf{X} \in \mathcal{A}$. This result implies that the linear programming relaxation

$$\underset{\mathbf{X}}{\text{maximize}} \quad \mathrm{Tr}(\mathbf{P}^T\mathbf{SX}) \qquad (9a)$$

$$\text{subject to: } 0 \le \mathbf{X}_{ji} \le 1, \forall i \in \mathcal{I}, j \in \mathcal{J} \qquad (9b)$$

$$\sum_{j=1}^{J} \mathbf{X}_{ji} = 1, \forall i \in \mathcal{I} \qquad (9c)$$

$$\sum_{i=1}^{I} \mathbf{X}_{ji} = c_j, \forall j \in \mathcal{J} \qquad (9d)$$

is always exact [i.e., problems (7) and (9) are equivalent]. The situation is graphically illustrated in Figure 4, which shows the geometry of (9) in relation to the geometry of (7).

Since (9) is an LP, it follows that either an interior point method or the simplex method can be used for solving (7). When using an interior point method, one should be mindful of cases where there are multiple Boolean solutions with the same (optimal) objective value because the interior point algorithm may converge to the center of a polyhedral facet (instead of a vertex), yielding a noninteger solution. We actually need a basic solution of the LP [32], and advanced interior

point LP solvers include means of identifying such a solution, e.g., [33]. These subtleties are avoided altogether if one uses the simplex method or, better yet, the shortest augmenting path algorithm [30], which has favorable low-order polynomial complexity even in the worst case. If only a general interior point LP solver is available, then a random perturbation heuristic can be applied, see [2].

The overall algorithm for (4) is now clear: one starts from a suitable initialization and iterates between updating $\mathbf{S}$ and updating $\mathbf{X}$. For initialization, one can use regular VQ/k-means to come up with an initial $\mathbf{S}$ without regard to capacity constraints. The sessions can be ordered according to population, and excess papers can be moved to the next session in line to produce an initial feasible assignment. Updating can start from $\mathbf{X}$ or from $\mathbf{S}$, and continue as long as the cost is reduced. Finally, initialization does matter (and VQ/k-means is itself sensitive with respect to initialization), so the overall algorithm should be initialized from different starting points 10–30 times to get close to the best possible results. The solution with the smallest cost is then chosen as the final one. At this point, the reader might rightfully wonder how well this algorithm works in practice, compared to expert human assignment. To get a sense of the kind of results that can be expected, see "How Well Does This Work? The ICASSP 2009/ SPCOM TC Case Study."

GAUGING THE OPTIMALITY GAP: SEMIDEFINITE RELAXATION

Even though the capacitated k-means clustering problem in (4) is NP-hard, it is possible to efficiently obtain a nontrivial lower bound on its optimal value. Notice that a tight lower bound also serves as a nice exploratory tool, e.g., it can be used to evaluate the performance of the GLM/LBG-based approximation algorithm. In obtaining this lower bound, we first demonstrate that the capacitated k-means clustering problem in (4) can be cast as a QCQP. This is an important link because the literature on quadratic optimization is rich and the tools that have been developed in the field of quadratic optimization are well understood.

In particular, we show that the capacitated k-means clustering problem in (4) can be cast in a form that closely resembles the (in)famous quadratic assignment problem (QAP) [34], [35]. Unlike the classical QAP, however, ours is a semiassignment problem, due to the particular structure of our set of admissible assignment matrices $\mathcal{A}$. Nonetheless, many relaxation strategies that have been developed for the QAP can be applied in our context as well. The best convex relaxations known for QAP are based on SDR. We also apply an SDR method [22], [36]–[39] to our problem. It is worth noting that a different SDR approach to (unconstrained) k-means clustering was pursued in [19].

The main reason why the capacitated k-means clustering problem (4) can be cast as a QCQP is that the optimal $\mathbf{S}^\bullet$ can be analytically derived as a function of $\mathbf{X}$; that is, the cost function can be concentrated with respect to $\mathbf{S}$ for a given $\mathbf{X}$. There are no constraints on $\mathbf{S}$; therefore, the minimizer is given by $\mathbf{S}^\bullet = \mathbf{PX}^\dagger$, where $\mathbf{X}^\dagger$ denotes the Moore–Penrose pseudoinverse of $\mathbf{X}$. It

HOW WELL DOES THIS WORK? THE ICASSP 2009/SPCOM TC CASE STUDY

The list of accepted papers from the SPCOM TC track of ICASSP 2009 is used for validation. There were 132 papers accepted, which were to be split among a total of 14 sessions: six lectures and eight poster sessions, containing six and 12 papers each, respectively. The algorithmic results will be compared to the final technical program that was manually produced by Nicholas D. Sidiropoulos, who chaired SPCOM TC at the time.

The list of key words (features) was manually produced by the authors, parsing the list of paper titles. Each title was examined, existing key words were added to the paper as appropriate, and new key words were created and added to list of key words as needed. The final list contains a total of 44 key words:

optimization, cross-layer, networking, resource, QCSI, game, precoding, DSL, distributed, sensor, sparse, MIMO, detection, performance, blind, cognitive, cooperative, capacity, network, coding, security, multiuser, beamforming, downlink, relay, uplink, CDMA, OFDM, synchronization, turbo, quantization, equalization, interference, estimation, training, tracking, localization, consensus, diversity, PAR, STBC, FH, scheduling, communications.

The feature vector of each paper is 44×1, with ones in the positions corresponding to features it possesses, and zeros elsewhere. The median number of (nonzero) features per paper was three.

The computer-generated conference program (using the algorithm in the section "Proposed Algorithm for Paper-to-Session Assignment") for ICASSP 2009/SPCOM TC is listed as Appendix A (available as supplementary material accompanying this article in IEEE *Xplore*). Session pseudotitles were produced by session centroid thresholding. If a key word is included in more than 30% of the papers in a session (the corresponding centroid element is greater than 0.3), then the key word is included in the session pseudotitle. Note that the order of key words in the pseudotitles is arbitrary (one could list them in order of importance, determined by the magnitude of centroid elements). The listed computer-generated program attains a (sum-of-squares) cost of 148.1 (after 30 initializations). The actual technical program that was manually produced by Sidiropoulos attains a cost of 187.25, primarily because, after two days of manual optimization and with a looming deadline ahead, he gave up and used an "umbrella" poster session for papers that did not fit elsewhere but otherwise had little in common. This is avoided in the solution listed in the supplementary material (Appendix A) available in IEEE *Xplore*, and in several other suboptimal solutions, which typically have a few discrepancies but avoid umbrella sessions. Note also that the running time of the algorithm in the section "Proposed Algorithm for Paper-to-Session Assignment" was less than 1.5 minutes (on a Dell E64000 laptop) for this data set, for 30 runs from different initial points.

follows that the conference program optimization problem in (4) can be written equivalently as

$$\underset{X}{\text{minimize}} \quad \| P - PX^{\dagger}X \|_F^2 \tag{10a}$$

$$\text{subject to: } X \in \mathcal{A}. \tag{10b}$$

Since any $X \in \mathcal{A}$ is full row rank, the pseudoinverse has the form $X^{\dagger} = X^T (XX^T)^{-1}$. Using the property (5), this reduces to the simpler form $X^{\dagger} = X^T \Lambda^{-1}$. It follows that (10) is equivalent to the problem

$$\underset{X}{\text{minimize}} \quad \| P - PX^T \Lambda^{-1} X \|_F^2 \tag{11a}$$

$$\text{subject to: } X \in \mathcal{A}. \tag{11b}$$

Expanding the squares in the objective of (11), we get $\| P - PX^T \Lambda^{-1} X \|_F^2 =$

$$= \| P \|_F^2 - 2\text{Tr}(P^T PX^T \Lambda^{-1} X) + \| PX^T \Lambda^{-1} X \|_F^2$$
$$= \| P \|_F^2 - 2\text{Tr}(P^T PX^T \Lambda^{-1} X) + \| PX^T \Lambda^{-1/2} \|_F^2$$
$$= \| P \|_F^2 - \| PX^T \Lambda^{-1/2} \|_F^2,$$

where we have used $\| Y \|_F^2 = \text{Tr}(YY^T)$, so

$$\| PX^T \Lambda^{-1} X \|_F^2 = \text{Tr}(PX^T \Lambda^{-1} XX^T \Lambda^{-1} XP^T),$$

and, since in this particular case, $X^T \Lambda^{-1} = X^{\dagger}$,

$$\| PX^T \Lambda^{-1} X \|_F^2 = \text{Tr}(PX^T \Lambda^{-1} XP^T)$$
$$= \text{Tr}(PX^T \Lambda^{-1/2} \Lambda^{-1/2} XP^T) = \| PX^T \Lambda^{-1/2} \|_F^2.$$

Hence, the problem in (11) can be expressed equivalently as

$$\underset{X}{\text{minimize}} \quad \| P \|_F^2 - \| PX^T \Lambda^{-1/2} \|_F^2 \tag{12a}$$

$$\text{subject to: } X \in \mathcal{A}. \tag{12b}$$

This is now a quadratic minimization problem subject to Boolean constraints, which is intractable [formally, the NP-hardness of (12) follows from its equivalence to (4)]. The form (12) closely resembles the QAP: The difference is that, in (12), X is constrained to lie in $\mathcal{A}$ instead of the set of permutation matrices, as in the classical QAP.

To illustrate how one can apply SDR to the problem above, we write it first in a more clear form using simple algebraic manipulations. Problem (12) can be written equivalently as

$$\underset{X}{\text{minimize}} \quad \| P \|_F^2 - \| (P \otimes \Lambda^{-1/2}) \text{vec}(X) \|_2^2 \tag{13a}$$

$$\text{subject to: } X \in \mathcal{A}, \tag{13b}$$

where $\otimes$ denotes the Kronecker product operation, and vec the operator that stacks the columns of a matrix into one vector. Recall that the linear system of equations (3b) and (3c) can be written in the form $Gx = d$, where $x = \text{vec}(X)$ and define the matrices
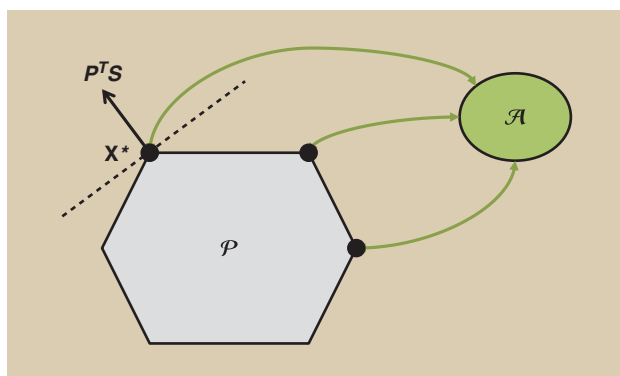
[FIG4] The feasible set of (9), which is a polyhedron, is shaded and denoted as $\mathcal{P}$. The objective $\mathrm{Tr}(\mathbf{P}^{\mathsf{T}}\mathbf{SX})$ is linear, and the point $\mathbf{X}^{\cdot}$ is optimal; it is the point in $\mathcal{P}$ as far as possible in the direction $\mathbf{P}^{\mathsf{T}}\mathbf{S}$. As illustrated in the figure, the polyhedron $\mathcal{P}$ is such that its vertices are all points in the set $\mathcal{A}$, and thus Boolean (see the corresponding definition of $\mathcal{A}$ in the text).

$$\mathbf{Q} \triangleq \mathbf{P} \otimes \mathbf{\Lambda}^{-1/2}, \text{ and } \mathbf{L} \triangleq \begin{bmatrix} \mathbf{G}^{\mathsf{T}}\mathbf{G} & -\mathbf{G}^{\mathsf{T}}\mathbf{d} \\ -\mathbf{d}^{\mathsf{T}}\mathbf{G} & \mathbf{d}^{\mathsf{T}}\mathbf{d} \end{bmatrix}.$$

With these definitions, (13) can be written equivalently as

$$\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{P}\|_F^2 - \mathbf{x}^{\mathsf{T}}\mathbf{Q}^{\mathsf{T}}\mathbf{Q}\mathbf{x} \qquad (14\text{a})$$

$$\text{subject to:} \quad \text{diag}(\mathbf{x}\mathbf{x}^{\mathsf{T}}) = \mathbf{x}, \qquad (14\text{b})$$

$$\|\mathbf{G}\mathbf{x} - \mathbf{d}\|_2^2 = 0. \qquad (14\text{c})$$

This is now a standard form QCQP, the quadratic constraints in (14b) ensuring that all variables $\mathbf{x}_i$ are Boolean. Let us illustrate how one can apply SDR to the above problem step by step.

TECHNICAL DETAILS OF SDR
Using the fact that $\mathbf{x}^{\mathsf{T}}\mathbf{Q}^{\mathsf{T}}\mathbf{Q}\mathbf{x} = \mathrm{Tr}(\mathbf{x}^{\mathsf{T}}\mathbf{Q}^{\mathsf{T}}\mathbf{Q}\mathbf{x}) = \mathrm{Tr}(\mathbf{Q}^{\mathsf{T}}\mathbf{Q}\mathbf{x}\mathbf{x}^{\mathsf{T}})$ and the change of variables

$$\mathbf{W} = \begin{bmatrix} \mathbf{x}\mathbf{x}^{\mathsf{T}} & \mathbf{x} \\ \mathbf{x}^{\mathsf{T}} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{1,1} & \mathbf{W}_{1,2} \\ \mathbf{W}_{1,2}^{\mathsf{T}} & \mathbf{W}_{2,2} \end{bmatrix}, \qquad (15)$$

problem (14) is reformulated in a higher dimensional space as follows:

$$\underset{\mathbf{W}}{\text{minimize}} \quad \|\mathbf{P}\|_F^2 - \mathrm{Tr}(\mathbf{W}_{1,1}\mathbf{Q}^{\mathsf{T}}\mathbf{Q}) \qquad (16\text{a})$$

$$\text{subject to:} \quad \text{diag}(\mathbf{W}_{1,1}) = \mathbf{W}_{1,2}, \; \mathbf{W}_{2,2} = 1, \qquad (16\text{b})$$

$$\mathrm{Tr}(\mathbf{L}\mathbf{W}) = 0, \qquad (16\text{c})$$

$$\mathbf{W} \succeq 0, \; \mathrm{rank}(\mathbf{W}) = 1. \qquad (16\text{d})$$

Here, $\mathbf{W}_{1,1}$ denotes the $JI \times JI$ upper-left block, $\mathbf{W}_{1,2}$ the $JI \times 1$ upper-right block, and $\mathbf{W}_{2,2}$ the 1×1 lower-right block of the $(JI + 1) \times (JI + 1)$ matrix $\mathbf{W}$. Problem (16) is equivalent to (14), since any rank-1 matrix satisfying (16b) can be factored according to the definition in (15), and, hence, the solution of (14) can be easily constructed from the solution of (16) and vice versa. The only difficult part of (16) is the nonconvex rank-1 constraint on $\mathbf{W}$. Dropping this constraint yields an SDR of (16)

$$\underset{\mathbf{W}}{\text{minimize}} \quad \|\mathbf{P}\|_F^2 - \mathrm{Tr}(\mathbf{W}_{1,1}\mathbf{Q}^{\mathsf{T}}\mathbf{Q}) \qquad (17\text{a})$$

$$\text{subject to:} \quad \text{diag}(\mathbf{W}_{1,1}) = \mathbf{W}_{1,2}, \; \mathbf{W}_{2,2} = 1, \qquad (17\text{b})$$

$$\mathrm{Tr}(\mathbf{L}\mathbf{W}) = 0, \qquad (17\text{c})$$

$$\mathbf{W} \succeq 0. \qquad (17\text{d})$$

In contrast with (16), problem (17) is convex (in fact, a semidefinite program), and it can be readily solved in polynomial time using efficient interior point methods [18]. If the solution $\mathbf{W}^{\cdot}$ of this semidefinite program turns out to have rank 1, then it is a solution for (16) as well. However, because of the relaxation, $\mathbf{W}^{\cdot}$ will not always be a rank-1 matrix; hence, the optimal value of (17) generally provides a lower bound on the optimal value of (16) [note that (4) and (16) have the same optimal value].

Given $\mathbf{W}^{\cdot}$, an approximate solution for the technical program optimization problem in (4) can be produced using a procedure known as *Gaussian randomization* [37]. This procedure consists of three main steps: 1) draw a random vector $\mathbf{v} = [v_1, \cdots, v_{JI+1}]^{\mathsf{T}}$ from $\mathcal{N}(0, \mathbf{W}^{\cdot})$, 2) form the new vector ξ consisting of the first JI entries of $\mathbf{v}$ divided by v_{JI+1}, and 3) find the vector that is closest to ξ and is feasible for (14), i.e., the vector $\mathbf{x}$ that minimizes $\|\xi - \mathbf{x}\|_2^2$ subject to (14b)–(14c).

This three-step procedure can be repeated a number of times, and the vector that gives the smallest objective value in (14) can be eventually chosen as an approximate solution. The intuition behind randomization is that it will generate candidate solutions that are close to the eigenvector of $\mathbf{W}^{\cdot}$ that corresponds to the largest eigenvalue, but will also take the other eigenvalues into account when these are large enough. Randomization has been widely used in the quadratic optimization literature, and its merits are well documented; see [37, Section IV] for an excellent discussion on this issue.

The rounding problem in step 3) seems hard, but it is not. To explain this, note that for any $\mathbf{x}$ feasible for (14), we have that $\|\xi - \mathbf{x}\|_2^2 = \|\xi\|_2^2 - \mathbf{x}^{\mathsf{T}}\xi + I$, and therefore, rounding corresponds to

$$\underset{\mathbf{x}}{\text{maximize}} \quad \mathbf{x}^{\mathsf{T}}\xi \text{ subject to: (14b)–(14c).} \qquad (18)$$

Notice that the constraint in (14c) is equivalent to the convex constraint $\mathbf{G}\mathbf{x} = \mathbf{d}$, and, since $\mathbf{G}$ is a totally unimodular matrix, problem (18) can be solved efficiently in polynomial time, using, e.g., the shortest augmenting path algorithm from [30]. The same discussion as that for problem (7) applies for (18) as well.

COMPLEXITY CONSIDERATIONS
It is important to recognize that the alternating optimization algorithm in the section "Proposed Algorithm for Paper-to-Session Assignment" is much cheaper and faster than the SDR approach in the section "Gauging the Optimality Gap: Semidefinite Relaxation." This is similar to classical k-means, and it is the reason why alternating optimization is so popular in applications of k-means clustering. For alternating optimization, the conditional update of $\mathbf{S}$ is very simple; the most expensive part in every iteration is the conditional update of $\mathbf{X}$. The shortest augmenting path algorithm from [30] can carry out the $\mathbf{X}$-update in time $O(JI^2)$. Linear programming (either with an interior point or with a simplex method) can be effectively used for the $\mathbf{X}$-update as well. In relation to the alternating optimization algorithm of the section "Proposed Algorithm

for Paper-to-Session Assignment," the computational disadvantage of SDR in the section "Gauging the Optimality Gap: Semidefinite Relaxation" stems from the fact that it lifts the problem in a higher dimensional space [in (15)–(16)], and this lifting squares the number of variables. This implies much higher complexity. Two important advantages of the SDR approach, on the other hand, are that it yields an approximation in one shot (read: with a predictable number of interior-point iterations for the relaxed convex problem), and it also yields a bound on how far any solution is from an optimum one. The latter is something that cannot be gauged from alternating optimization.

VARIATIONS OF THE BASIC FORMULATION

There are several variations of the basic formulation that one can readily envision. We now briefly mention a few interesting alternatives.

WEIGHTING

In some cases, the TPC chair may wish to highlight emerging or important areas in the technical program. This can be accomplished via feature weighting, i.e., optimizing a weighted least squares cost of the form

$$\| \mathbf{D}(\mathbf{P} - \mathbf{S}\mathbf{X}) \|_F^2,$$

where $\mathbf{D}$ is a full-rank diagonal matrix holding the feature weights. Such weighting can be absorbed in $\mathbf{P}$ and $\mathbf{S}$, and, since the latter is unconstrained, it does not change the essence of the proposed solutions. It is clear that the proposed GLM/LBG algorithm can be readily modified to handle this extension. Following steps similar to (10)–(12), it is a simple exercise to verify that the SDR approach can be extended as well.

ALIGNMENT WITH ORGANIZATIONAL STRUCTURE

Organizations such as the NSF often prefer to form panels that reflect their organizational structure. For example, for a large cross-disciplinary solicitation that falls under the auspices of multiple divisions (sometimes even across directorates), from a logistics point of view, it makes a lot of sense to produce panels that are reasonably well aligned with the constituent programs. This can be accomplished by anchoring panel centroids in $\mathbf{S}$ not to deviate too far from the constituent organizational unit profiles, stored in $\mathbf{S}_o$, i.e., by augmenting the cost function in (4) with a penalty term as

$$\| \mathbf{P} - \mathbf{S}\mathbf{X} \|_F^2 + \rho \| \mathbf{S}_o - \mathbf{S} \|_F^2.$$

By varying the penalty parameter $\rho > 0$, one can trade off between alignment and homogeneity. Notice that this augmentation does not fundamentally change the nature of our solutions. In fact, the optimal session centroid matrix $\mathbf{S}^\star$ is still given in simple closed form as $\mathbf{S}^\star = (\mathbf{P}\mathbf{X}^T + \rho \mathbf{S}_o)(\mathbf{\Lambda} + \rho \mathbf{I})^{-1}$. As a result, both the alternating optimization algorithm and the proposed SDR approach can be easily modified to account for this penalty term.

DIVIDE-AND-CONQUER AND TREE-STRUCTURED VQ

For the special case where we are interested in splitting the papers into just $J = 2$ sessions, the conditional update of $\mathbf{X} = $

$[\mathbf{x}_1 \mathbf{x}_2]^T \in \mathbb{R}^{2 \times I}$ given $\mathbf{S} = [\mathbf{s}_1 \mathbf{s}_2] \in \mathbb{R}^{N \times 2}$ takes a very simple form. This simplification can be used to construct a divide-and-conquer algorithm for paper-to-session assignment, reminiscent of hierarchical clustering approaches and tree-structured VQ [20], [21]. Consider the conditional paper-to-session assignment problem for $J = 2$ sessions only. Using the equivalence shown in (6) and (7), the optimization problem is

$$\operatorname*{maximize}_{\mathbf{x}_1, \mathbf{x}_2} \ \mathrm{Tr}(\mathbf{P}^T [\mathbf{s}_1 \mathbf{s}_2] [\mathbf{x}_1 \mathbf{x}_2]^T), \tag{19a}$$

$$\text{subject to: } \mathbf{x}_1(i) \in \{0, 1\}, \mathbf{x}_2(i) \in \{0, 1\}, \forall i \in \mathcal{I}, \tag{19b}$$

$$\mathbf{x}_1(i) + \mathbf{x}_2(i) = 1, \forall i \in \mathcal{I}, \tag{19c}$$

$$\sum_{i=1}^{I} \mathbf{x}_1(i) = c_1, \sum_{i=1}^{I} \mathbf{x}_2(i) = c_2 = I - c_1. \tag{19d}$$

Using the constraints (19b)–(19d) and the fact that $\mathrm{Tr}(\mathbf{A}\mathbf{B}) = \mathrm{Tr}(\mathbf{B}\mathbf{A})$, one can eliminate variable $\mathbf{x}_2$ from (19), yielding the simpler problem

$$\operatorname*{maximize}_{\mathbf{x}_1} \ \mathbf{x}_1^T \mathbf{P}^T (\mathbf{s}_1 - \mathbf{s}_2)$$
$$\text{subject to: } \mathbf{x}_1(i) \in \{0, 1\}, \forall i \in \mathcal{I},$$
$$\sum_{i=1}^{I} \mathbf{x}_1(i) = c_1,$$

from which it is clear that the optimal solution is to allocate the c_1 units to the c_1 largest elements of $\mathbf{P}^T (\mathbf{s}_1 - \mathbf{s}_2)$. These can be found using a sorting operation, at complexity $O(I \log I)$, or by direct parsing at $O(I c_1)$.

Now, using the above result for $J = 2$, we can construct a potentially appealing divide-and-conquer solution for the paper-to-session assignment problem for $J > 2$ as follows: We start with regular VQ/k-means to produce an initial centroid matrix $\mathbf{S}$, the columns of which are then ordered according to paper population. In the divide step, we first process the columns of $\mathbf{S}$ (e.g., using plain 2-means) to produce two new (super)centroids, then use the sorting-based algorithm to assign papers to these two centroids in a way that respects the session capacity constraints. We then recursively refine and conquer the subproblems in a similar manner. Once we produce the final assignment, we update $\mathbf{S}$ and repeat the procedure. This algorithm is fast and can be quite effective, mainly depending on the quality of the initialization point.

THE REVIEW ASSIGNMENT PROBLEM

The review assignment stage is even more difficult than putting together the final technical program simply because it involves (a lot) more papers and every paper must be reviewed by more than one reviewer. Suppose that I papers are to be assigned for review to (at most) J reviewers. Reviewer j has a fixed vector profile $\mathbf{s}_j$ representing the reviewer's expertise and reviewing interests, and a prenegotiated reviewing capacity r_j. Every paper should be reviewed by, say, three reviewers. Our goal here is to minimize the paper-to-reviewer mismatches, i.e., a paper should be assigned for review to three reviewers whose individual vector profiles cover as much as

possible the paper profile $\mathbf{p}_i$. At the same time, and of equal importance, is that the reviewer profiles should collectively cover the paper profile $\mathbf{p}_i$ as much as possible.

One can thus pose the review assignment problem as follows:

$$\underset{\mathbf{X}\in\{0,1\}^{J\times 3I}}{\text{minimize}} \quad (1-\lambda)\sum_{i=1}^{3I}\mathbf{1}^T\Big(\mathbf{P}_{*\lceil\frac{i}{3}\rceil}-\mathbf{S}\mathbf{X}_{*i}\Big)_+$$

$$+ \lambda\sum_{i=1}^{I}\mathbf{1}^T[\mathbf{P}_{*i}-\mathbf{S}(\mathbf{X}_{*3i-2}+\mathbf{X}_{*3i-1}+\mathbf{X}_{*3i})]_+, \quad (21a)$$

$$\text{subject to:} \quad \mathbf{X}_{ij}\in\{0,1\}, \,\forall i,j, \quad (21b)$$

$$\sum_{j=1}^{J}\mathbf{X}_{ji}=1, \,\forall i\in\{1,\ldots,3I\}, \quad (21c)$$

$$\sum_{i=1}^{3I}\mathbf{X}_{ji}\le r_j, \,\forall j\in\{1,\ldots,J\}, \quad (21d)$$

$$\sum_{i=3k+1}^{3k+3}\mathbf{X}_{ji}\le 1, \,\forall j\in\{1,\ldots,J\},$$

$$\forall k\in\{0,\cdots,I-1\}, \quad (21e)$$

$$\mathbf{X}_{ij}=0 \quad \forall(i,j)\in COI. \quad (21f)$$

Here, $\mathbf{P}$ is the matrix of paper profiles, $\mathbf{S}=[\mathbf{s}_1,\cdots,\mathbf{s}_J]\in\mathbb{R}^{N\times J}$ is the matrix of the reviewer profiles, and $\mathbf{X}_{*i}$ denotes the ith column of $\mathbf{X}$. The symbol $(\cdot)_+$ denotes projection to the non-negative orthant, $\lceil\cdot\rceil$ denotes the ceiling function, and $\mathbf{1}$ denotes the $N\times 1$ vector of all ones.

Let us now explain the mathematical formulation of the review assignment problem in detail. Observe that the cost function in (21a) comprises two sums. The first aims to minimize the paper key words not covered by the associated reviewers individually, while the second (the collective span term) accounts for the paper key words that are not covered by the sum of profiles of the associated reviewers. The two cost factors are weighted using a suitable regularization parameter $0 < \lambda < 1$.

[FIG5] An illustration of the "dual" nature of the problem in terms of utility functions. Each paper has as utility its key word coverage (collectively, from all assigned reviewers), and each reviewer has as a utility the aggregate amount of key words matched from his/her assigned papers.

The inequality constraints in (21e) protect each paper from being assigned to the same reviewer twice, while the constraints (21c) and (21d) ensure that each paper $\mathbf{P}_{*i}$ will be assigned for review to three reviewers, while respecting the reviewer capacity constraints. In particular, columns $3i-2, 3i-1, 3i$ in $\mathbf{X}\in\{0,1\}^{J\times 3I}$ comprise Boolean variables, which select three different reviewers for paper i [see (21c) and (21e)]. Moreover, the ceiling operation $\lceil i/3 \rceil$ repeats three times the ith column of $\mathbf{P}$ (paper i) to calculate its mismatch with each of the three individual assigned reviewers [see (21a)].

Finally, note that reviewers should not have a conflict of interest (COI) with the papers they are reviewing (e.g., they cannot be from the same department as any of the paper's authors). In case there is a COI between a reviewer and specific papers, additional COI constraints must be included in the optimization. These are taken into account by the constraint in (21f), which enforces the pertinent assignment variables to be equal to zero.

The review assignment problem as posed in (21) is combinatorial, but it has a convex objective function, and also the constraints in (21c)–(21f) are convex constraints. Interestingly, replacing the Boolean constraints in (21b) by the convex inequality constraints $0\le\mathbf{X}_{ji}\le 1$ leads to a relaxation problem whose feasible set is a polyhedron with Boolean vertices only (we shall call this the *review assignment polyhedron*). This can be seen by noting that the coefficient matrix of the set of linear inequalities (21c)–(21f) is totally unimodular (see, e.g., [31]). Even so, problem (21) is difficult to solve due to the collective span term in the objective, which is a nonlinear function of $\mathbf{X}$. One can construct, however, an approximate solution through convex relaxation and rounding.

Before we explain this approach in detail, let us first discuss several interesting points that can be gauged from the problem formulation in (21). To simplify exposition and better highlight these points, we temporarily confine attention to the case of Boolean matrices $\mathbf{P}$ and $\mathbf{S}$. We emphasize, however, that the convex relaxation approach that we propose for (21) holds for general matrices $\mathbf{P}$ and $\mathbf{S}$.

REMARK 2: PAPER AND REVIEWER UTILITY FUNCTIONS

One may think of the review assignment problem in terms of utility functions. To see this, it is convenient to introduce some mathematical notation first. Suppose that both $\mathbf{P}$ and $\mathbf{S}$ are Boolean. Moreover, suppose that assignment $\mathbf{X}$ assigns paper $\mathbf{p}_i$ to the reviewer set $\mathcal{R}_i(\mathbf{X})$ (with $|\mathcal{R}_i(\mathbf{X})|=3$) and the same assignment $\mathbf{X}$ assigns to reviewer j the paper set $\mathcal{N}_j$ (with $|\mathcal{N}_j(\mathbf{X})|\le r_j$). Let $u_p^i(\mathbf{X})=\mathbf{1}^T\mathbf{p}_i-\mathbf{1}^T\big(\mathbf{p}_i-\sum_{k\in\mathcal{R}_i(\mathbf{X})}\mathbf{s}_k\big)_+$ be the utility function of paper i (in case of Boolean $\mathbf{P}$ and $\mathbf{S}$ this is paper i's collective key word coverage resulting from assignment $\mathbf{X}$), and let $u_r^j(\mathbf{X})=\sum_{k\in\mathcal{N}_j(\mathbf{X})}[\mathbf{1}^T\mathbf{p}_k-\mathbf{1}^T(\mathbf{p}_k-\mathbf{s}_j)_+]$ be the utility function of reviewer j (in case of Boolean $\mathbf{P}$ and $\mathbf{S}$, this is the total number of key word matches between the reviewer and all papers assigned to the reviewer). Maximizing reviewer satisfaction and paper utility can be conflicting objectives, as illustrated in Figure 5 and exemplified in Figure 6. The tradeoff between the two is captured in the problem formulation (21) because the objective function in (21a) can be written in terms of the $\{u_p^i(\mathbf{X})\}_{i=1}^I$ and $\{u_r^j(\mathbf{X})\}_{j=1}^J$, by regrouping terms accordingly. ∎

REMARK 3

Observe that for Boolean matrices $\mathbf{P}$ and $\mathbf{S}$ the first sum term in (21a) can be replaced by a function linear in $\mathbf{X}$ since, for any feasible assignment $\mathbf{X}$ and Boolean matrices $\mathbf{P}$ and $\mathbf{S}$, it holds that $\sum_{i=1}^{3I} \mathbf{1}^{\mathrm{T}}(\mathbf{P}_{*[(i/3)]} - \mathbf{S}\mathbf{X}_{*i})_{+} = \sum_{i=1}^{3I} \mathbf{1}^{\mathrm{T}}\mathbf{P}_{*[(i/3)]} - \sum_{i=1}^{3I} \mathbf{P}^{\mathrm{T}}_{*[(i/3)]} \ \mathbf{S}\mathbf{X}_{*i}$. In other words, this sum attempts to maximize the total affinity between papers and reviewers, which is reminiscent of the approach followed in [2]. ∎

Let us now turn the discussion to general $\mathbf{P}$ and $\mathbf{S}$, and describe explicitly the convex relaxation approach for (21). Let $\mathbf{X}^{\cdot}$ denote the solution to the relaxation program where the Boolean constraints $\mathbf{X}_{ji} \in \{0, 1\}$ are replaced by the interval ones, $0 \le \mathbf{X}_{ji} \le 1$. This relaxation yields a convex problem, which can be reformulated as an LP and solved efficiently. To see this, introduce for every individual summand in (21a) an associated slack variable t_i, and note that $\max(x, 0) \le t_i \Longleftrightarrow x \le t_i$ for $t_i \ge 0$. The constraint $\max(x, 0) \le t_i$ will always be satisfied with equality at the optimum, which yields the LP reformulation.

Unfortunately, however, $\mathbf{X}^{\cdot}$ is not guaranteed to be Boolean (the LP emerging after introducing the slack variables is not guaranteed to be totally unimodular); therefore, we need a way of converting the solution of the relaxed program into a good admissible solution for (21). This can be done by finding an assignment $\mathbf{X}$, which is as close as possible (in a Euclidean sense) to $\mathbf{X}^{\cdot}$, i.e., by finding an $\mathbf{X}$ that minimizes $\|\mathbf{X} - \mathbf{X}^{\cdot}\|^2_F$ subject to (21b)–(21f). This rounding problem seems hard, but it is not. To explain this, note that for any assignment $\mathbf{X}$ feasible for (21), we have that $\|\mathbf{X} - \mathbf{X}^{\cdot}\|^2_F = \|\mathbf{X}^{\cdot}\|^2_F - \mathrm{Tr}(\mathbf{X}^{\mathrm{T}}\mathbf{X}^{\cdot}) + 3I$, and, therefore, rounding corresponds to

$$\text{maximize } \mathrm{Tr}(\mathbf{X}^{\mathrm{T}}\mathbf{X}^{\cdot}) \text{ subject to (21b)–(21f).}$$

The above problem is equivalent to its linear programming relaxation (and is therefore easy to solve), since the polyhedron arising from the relaxation has only Boolean vertices [which are precisely the feasible set (21b)–(21f)]. To appreciate how well the proposed review assignment method works, see "How Well Does Automated Review Assignment Work? A SPAWC 2010 Case Study" and "Quantitative Assessment of Review Assignment Quality."

SOME VARIATIONS OF THE BASIC FORMULATION

ALTERNATIVE COST FUNCTIONS

For simplicity, we use the sum of inconsistencies in the cost of our formulation in (21). An interesting alternative would be to employ the sum of squares of inconsistencies, essentially putting more emphasis (and penalizing more) the bad assignments. Note that using the sum of squares of inconsistencies would still lead to a convex cost function.

CONTROLLING THE WORST MATCHING

It is possible to design the assignment while explicitly imposing an upper bound $T^{\cdot}$ on the cost of the worst paper-reviewer matching

$$\mathbf{1}^{\mathrm{T}}[\mathbf{P}_{*i} - \mathbf{S}(\mathbf{X}_{*3i-2} + \mathbf{X}_{*3i-1} + \mathbf{X}_{*3i})]_{+} \le T^{\cdot}, \ \forall i,$$
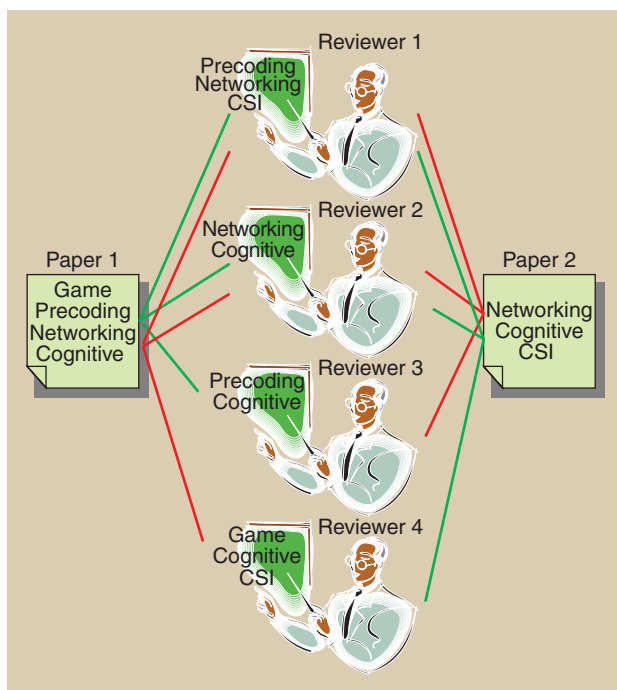
[FIG6] A green assignment: paper #1 utility = 3 + paper #2 utility = 3 ⇒ total paper utility = 6; reviewer utilities 2+2+2 (for paper #1) + 2+2+2 (for paper #2) ⇒ total reviewer utility = 12. Red assignment: paper #1 utility = 4 + paper #2 utility = 3 ⇒ total paper utility = 7; reviewer utilities 2+2+2 (for paper #1) + 2+2+1 (for paper #2) ⇒ total reviewer utility = 11. So, the green assignment is better in terms of reviewer utility, but the red one is better in terms of paper utility. This explains why the two objectives can be conflicting.

in addition to (21b)–(21e). This imposes a stricter requirement but changes the nature of the feasible set, as for general (even Boolean) $\mathbf{P}$ and $\mathbf{S}$, the new polyhedron is not guaranteed to have only Boolean vertices.

A more flexible approach to this issue is to consider varying λ in the cost function of (21) to trade off reviewer satisfaction for paper key word coverage. One can easily check the quality of a particular assignment after the optimization, by producing statistics, most notably how many key words of each paper have been collectively covered by its respective reviewers. If the result is not satisfactory, one can resolve the problem by changing λ so as to strike a more appropriate tradeoff. In fact, one can associate a different parameter $\lambda_i > 0$ to each paper i, if that is desired.

CONCLUSIONS

WHAT WE LEARNED

By viewing papers as vectors in a suitable feature space, the loosely defined tasks of paper-to-session and paper-to-reviewer assignment have been formulated as optimization problems that are strikingly familiar in many ways. The core problem underlying paper-to-session assignment is capacitated k-means, i.e., clustering under capacity constraints, and is NP-hard. For paper-to-reviewer assignment, it was shown that ensuring scientifically sound reviews (each aspect of each paper covered by at least one assigned reviewer) and

HOW WELL DOES AUTOMATED REVIEW ASSIGNMENT WORK? A SPAWC 2010 CASE STUDY

The submitted paper list and reviewing pool of SPAWC 2010 was used for validation. There were 203 submitted papers, and the reviewing pool comprised 64 reviewers (20+2+42 reviewers of capacity 8|15|16 papers, respectively). The list of key words (features) was manually produced by the authors by updating the previous list for ICASSP 2009; the final SPAWC key word list contained a total of 50 key words:

beamforming, blind, capacity, CDMA, classification, coding, cognitive, consensus, cooperative, cross-layer, detection, distributed, diversity, downlink, UWB, DSL, equalization, estimation, feedback, FH, game, joint source-channel, localization, MIMO, multiuser, network coding, networking, OFDM, optimization, par, performance, QCSI, quantization, random matrix, relay, resource, RFID, scheduling, security, sensor, sparse, speech-image, STBC, synchronization, time-varying, tracking, training, turbo, underwater, uplink.

The feature vector of each paper and each reviewer is 50×1, with ones in the positions corresponding to features it possesses, and zeros elsewhere. Feature vectors for the reviewers were created by Nicholas D. Sidiropoulos (acting as TPC chair), using his knowledge of their expertise. Feature vectors for the papers were partially entered by the respective authors, using a separate key word-clicking system that was set up for this purpose; however, not all authors obliged, so features for papers were also entered by Sidiropoulos after looking at paper titles. Parameter λ in the algorithm was set to $\lambda = 0.5$. It is worth mentioning that the ratio between the objective value of the linear programming relaxation and that of the rounded final solution was 98.7% [hinting that the final assignment was (at least) close to the optimal one]. The running time of the algorithm (relaxation + rounding) was less than two minutes for this data set, on a Dell E64000 laptop. The computer-generated review assignment is listed as Appendix B in the supplementary document accompanying this article in IEEE *Xplore*. Perusing this assignment, one can observe that four out of 64 reviewers were not assigned any paper at all in the final solution. In the cases where we have spare total reviewing capacity, we may consider adding a penalty term to avoid fully loading some reviewers and idling others.

maximizing reviewer satisfaction can be (and often are) conflicting objectives that must be traded against each other. The resulting paper-to-reviewer assignment problem is generally hard (albeit reducing it to a known NP-hard problem is not straightforward). Still, it was shown that it is possible to generate good suboptimal solutions using familiar signal processing tools. While there is certainly a lot more work to be done (e.g., automatic key word retrieval and paper mark-up, exploration of alternative problem formulations), our results indicate that computer-generated technical programs outperform expert manual work at a fraction of the time and with very limited input by the chair.

WHY IT IS IMPORTANT?

If you are a TPC chair, spend some time to come up with the right set of key words that capture what is happening in your area, invite enough good reviewers (a margin of 20% more reviewing capacity is always helpful, so do secure a few more reviewers; if you do not need all that reviewing power, reduce everyone's quota—they will be thankful). We have tested our algorithms with actual conference data, producing review and program assignments that TPC chairs have found very useful. We will make our algorithms freely available to the research community at the time of publication of this article.

As a final note, one can envision many other interesting applications of clustering under capacity constraints:

- assigning students to classrooms or study groups according to educational background, level of accomplishment in math/science/language, interests, etc.
- production-line packaging according to product quality features (e.g., tolerances)
- design of stock performance indices based on market sector, segment, capitalization, exposure to commodity price fluctuations, etc.

QUANTITATIVE ASSESSMENT OF REVIEW ASSIGNMENT QUALITY

We now discuss various performance metrics and statistics to appreciate the quality of the computer-generated solution.

DEFINITION: We define the quality index (QI) of a particular reviewer, as the average percentage of key word matches between the reviewer's profile and his/her assigned papers. As an example, suppose that a certain reviewer is assigned two papers for review, the papers having five and six key words, respectively, and let us assume that there are two key word matches from the first paper and three matches from the second. The reviewer's QI is then calculated as the average $((2/5 + 3/6)/2) \times 100\% = 45\%$.

The reviewers' QIs for the SPAWC 2010 case study can be found in the supplementary material (Appendix B), together with the optimized assignment. One can observe that 39/60 utilized reviewers had a QI above 80%, 54/60 reviewers had a QI above 70%, and all 60 utilized reviewers had a QI above 40%. From the collective span point of view, note that 187/203 papers ($\approx 92\%$) were fully covered (collectively) by their respective reviewers; the few papers that were not fully covered are marked with an asterisk in Appendix B.

As a final measure of the quality of the overall assignment, we compute the percentage of the overall key word matches, i.e., the total number of paper key words covered collectively by all assigned reviews. The percentage ratio (covered key words/total key words) was 98.1% for the SPAWC 2010 computerized assignment, indicating the high quality of the solution.

FURTHER INFORMATION

This article has supplementary downloadable material available in IEEE *Xplore*; see http://ieeexplore.ieee.org. The material includes a computer-generated conference program and a computer-generated review assignment using the methods presented in this article. Contact nicos@umn.edu for further questions regarding this work. In addition, a companion Web site is under development, and a link will be posted at http://www.ece.umn.edu/~nikos/.

AUTHORS

Nicholas D. Sidiropoulos (nikos@umn.edu) received his diploma in electrical engineering from the Aristotelian University of Thessaloniki, Greece, and his M.S. and Ph.D. degrees in electrical engineering from the University of Maryland in 1988, 1990, and 1992, respectively. He has been a professor at the University of Minnesota since 2011. His current research focuses primarily on signal and tensor analytics with applications in cognitive radio and big data. He received the National Science Foundation/ CAREER Award in 1998 and the IEEE Signal Processing Society (SPS) Best Paper Award in 2001, 2007, and 2011. He was an IEEE SPS Distinguished Lecturer (2008–2009) and chair of the IEEE Signal Processing for Communications and Networking Technical Committee (2007–2008). He was an associate editor of *IEEE Transactions on Signal Processing* (2000–2006), *IEEE Signal Processing Letters* (2000–2002), and on the editorial board of *IEEE Signal Processing Magazine* (2009–2011). He has been an area editor of *IEEE Transactions on Signal Processing* since 2012. He received the 2010 IEEE SPS Meritorious Service Award. He is a Fellow of the IEEE.

Efthymios E. Tsakonas (tsakonas@kth.se) received his diploma degree in electronics and computer engineering from the Technical University of Crete, Greece, 2008. Since September 2009, he has been a member of the Signal Processing Laboratory, Royal Institute of Technology (KTH), Stockholm, Sweden, where he earned his Ph.D. degree in 2014. His research interests are in convex optimization and approximation theory and algorithms, with applications in signal processing, communications, and preference measurement.

REFERENCES

[1] D. Hartvigsen, J. Wei, and R. Czuchlewski, "The conference paper-reviewer assignment problem," *Decision Sci.*, vol. 30, no. 3, pp. 865–876, 1999.

[2] C. Taylor, "On the optimal assignment of conference papers to reviewers," Tech. Rep. MS-CIS-08-30, Dept. Comput. Inform. Sci., Univ. Pennsylvania, Dec. 2008.

[3] S. Hettich and M. Pazzani, "Mining for proposal reviewers: Lessons learned at the National Science Foundation," in *Proc. 12th Association for Computing Machinery's Special Interest Group on Knowledge Discovery and Data Mining (ACM SIGKDD)*, 2006, pp. 862–871.

[4] M. Karimzadehgan, C. Zhai, and G. Belford, "Multi-aspect expertise matching for review assignment," in *Proc. 17th ACM Conf. Information and Knowledge Management (CIKM 2008)*, Napa Valley, CA, pp. 1113–1122.

[5] S. Theodoridis and K. Koutroumbas, *Pattern Recognition,* 4th Ed. New York: Academic Press, 2008.

[6] C. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

[7] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*. Boston, MA: ACM Press and Addison-Wesley, 1999.

[8] E. Papalexakis, N. D. Sidiropoulos, and M. Garofalakis, "Reviewer profiling using sparse matrix regression," in *Proc. 2010 IEEE International Conf. Data Mining (ICDM) Workshops/Workshop on Optimization Based Methods for Emerging Data Mining Problems (OEDM)*, Sydney, Australia, Dec. 2010, pp. 1214–1219.

[9] S. Dumais and J. Nielsen, "Automating the assignments of submitted manuscripts to reviewers," in *Proc. Association for Computing Machinery Special Interest Group on Information Retrieval*, 1992, pp. 233–244.

[10] F. Wang, B. Chen, and Z. Miao, "A survey on reviewer assignment problem," *Lect. Notes Comput. Sci.,* vol. 5027, pp. 718–727, 2008.

[11] S. Janak, M. Taylor, C. Floudas, M. Burka, and T. Mountziaris. (2006). A novel and effective integer optimization approach for the NSF panel assignment problem: A multi-resource and preference-constrained generalized assignment problem. *Ind. Eng. Chem*. Res. 45, pp. 258–265. [Online]. Available: http://ares.princeton.edu/casl_nsf_mgap/login.php

[12] P. Hahn, B. Kim, M. Guignard, J. Smith, and Y. Zhu, "An algorithm for the generalized quadratic assignment problem," *Comput. Optim. Appl.* (Springer), vol. 40, no. 3, pp. 351–372, July 2008.

[13] R. Burkard, M. Dell'Amico, and S. Martello, *Assignment Problems*. Philadelphia, PA: SIAM, 2009.

[14] A. Jain, M. Murty, and P. Flynn, "Data clustering: A review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, Sept. 1999.

[15] K. Wagstaff, C. Cardie, S. Rogers, and S. Schroedl, "Constrained k-means clustering with background knowledge," in *Proc. 18th Int. Conf. Machine Learning*, 2001, pp. 577–584.

[16] I. Davidson and S. Ravi, "Clustering with constraints: Feasibility issues and the k-means algorithm," in *Proc. 2005 SIAM Int. Conf. Data Mining*.

[17] P. Bradley, K. Bennett, and A. Demiriz, "Constrained k-means clustering," Tech. Rep. TR-2000-65, Microsoft Research, 2000.

[18] S. Boyd and L. Vandenberghe. (2004). *Convex Optimization*. [Online]. Cambridge, U.K.: Cambridge Univ. Press. Available: http://www.stanford.edu/~boyd/cvxbook.html

[19] J. Peng and Y. Wei, "Approximating k-means-type clustering via semidefinite programming," *SIAM J. Optim.*, vol. 18, no. 1, pp. 186–205, 2007.

[20] K. Jain and R. Dubes, *Algorithms for Clustering Data*. Englewood Cliffs, NJ: Prentice Hall, 1988.

[21] L. Kaufman and P. Rousseeuw, *Finding Groups in Data, an Introduction to Cluster Analysis*. New York: Wiley, 1990.

[22] H. Wolkowicz, "Relaxations of Q2P," in *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, H. Wolkowicz, R. Saigal, and L. Vandenberghe, Eds. Norwell, MA: Kluwer, 2000, Ch. 13.4, pp. 395–419.

[23] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. 28, pp. 129–137, Mar. 1982.

[24] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inform. Theory*, vol. 6, pp. 7–12, Mar. 1960.

[25] Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. 28, no. 1, pp. 84–95, Jan. 1980.

[26] D. Aloise, A. Deshpande, P. Hansen, and P. Popat, "NP-hardness of Euclidean sum-of-squares clustering," *Mach. Learn.*, vol. 75, pp. 245–248, May 2009.

[27] S. Dasgupta and Y. Freund, "Random projection trees for vector quantization," *IEEE Trans. Inform. Theory*, vol. 55, pp. 3229–3242, July 2009.

[28] P. Crescenzi and V. Kann. A compendium of NP optimization problems. [Online]. Available: http://www.nada.kth.se/~viggo.

[29] N. Guttmann-Beck and R. Hassin, "Approximation algorithms for minimum sum p-clustering," *Disc. Appl. Math.* vol. 89, nos. 1–3, pp. 125–142, Dec. 1998.

[30] J. Kennington and Z. Wang, "A shortest augmenting path algorithm for the semi-assignment problem," *Operations Res.*, vol. 40, no. 1, pp. 178–187, Feb. 1992.

[31] G. Birkhoff, "Tres observaciones sobre el algebra lineal," *Universidad Nacional de Tucuman Revista*, vol. 5, pp. 147–151, 1946.

[32] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*. Belmont, MA: Athena Scientific, 1997.

[33] MOSEK ApS. The MOSEK optimization toolbox for MATLAB manual. [Online]. Available: http://docs.mosek.com/7.0/toolbox.pdf

[34] Q. Zhao, S. Karisch, F. Rendl, and H. Wolkowicz, "Semidefinite programming relaxations for the quadratic assignment problem," *J. Combinatorial Optim.*, vol. 2, no. 1, pp. 71–109, 1998.

[35] J. Povh and F. Rendl, "Copositive and semidefinite relaxations of the quadratic assignment problem," *Discrete Optim.*, vol. 6, no. 3, pp. 231–241, Aug. 2009.

[36] M. Goemans and D. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problem using semidefinite programming," *J. ACM,* vol. 42, no. 6, pp. 1115–1145, Nov. 1995.

[37] Z. Q. Luo, W. K. Ma, A. M. C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[38] P. Tseng, "Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation," *SIAM J. Optim.*, vol. 14, no. 1, pp. 268–283, July 2003.

[39] S. Zhang, "Quadratic maximization and semidefinite relaxation," *Math. Program.*, vol. 87, no. 3, pp. 453–465, May 2000.

[40] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh, "Clustering with Bregman divergences," *J. Mach. Learn. Res.*, vol. 6, pp. 1705–1749, 2005.

[SP]

Fredrik Rusek, Giulio Colavolpe,
and Carl-Erik W. Sundberg

[lecture **NOTES**]

40 Years with the Ungerboeck Model: A Look at Its Potentialities

It has been about 40 years since Gottfried Ungerboeck published his paper [1] on an alternative maximum-likelihood (ML) detector for intersymbol interference (ISI) channels. The ISI model used by Ungerboeck is commonly referred to as the *Ungerboeck model*. Ungerboeck's ML detector has equivalent performance compared to Forney's detector, which was published two years earlier in [2], but received lesser considerations. Perhaps the best example of this is the fact that a BCJR algorithm [3] operating on the Ungerboeck model was derived as late as 2005 [4]. However, the Ungerboeck model has many strong aspects and has therefore been rediscovered over the last few decades.

SCOPE

In this lecture note, we give a number of illuminating examples where the Ungerboeck model is essential. We hope that this column will lead to increased awareness and use of the Ungerboeck model among the signal processing community.

RELEVANCE

Essentially all communication systems are modeled by a discrete-time model. The white-noise model is the predominant choice of model today. When low-complexity algorithms are used, the choice of model plays a role. In some cases, superior performance and/or lower complexity can be achieved by the very same algorithm, but where the white noise model has been replaced by another model. Awareness of models other than the white-noise model

is of great value to engineers and researchers, especially to those working in the borderline of signal processing and wireless communications.

PREREQUISITES

This lecture note assumes basic knowledge of signal space descriptions of communication systems, about Viterbi- and Bahl-Cooke-Jelinek-Raviv (BCJR)-type algorithms for communication channels with memory, and about factor graphs (FGs) and the sum-product algorithm (SPA).

SUMMARY OF DETECTION THEORY RESULTS

In the following, we will denote by $p(\cdot)$ [respectively, $P(\cdot)$] the probability density function (pdf) [respectively, the probability mass function (pmf)] of a continuous (respectively, discrete) random variable. In addition $\mathcal{R}(\cdot)$ and $\mathcal{I}(\cdot)$ denote the real and imaginary part of a complex number whereas $(\cdot)^*$ and $(\cdot)^T$ stand for transpose conjugate and transpose, respectively. Let $x = [x_0, x_1, \cdots]$ be a sequence of modulation symbols drawn from a discrete alphabet $\mathcal{X}$. These symbols are transmitted over a communication channel via a modulation format. Let z be an arbitrary sufficient statistic properly extracted from the received signal. Maximum a posteriori (MAP) sequence and symbol detection strategies are based on the following decision rules:

$$\hat{x} = \underset{x}{\arg\max} P(x \mid z)$$
$$= \underset{x}{\arg\max}\, p(z \mid x) P(x) \qquad (1)$$

and

$$\hat{x}_k = \underset{x_k}{\arg\max} P(x_k \mid z)$$
$$= \underset{x_k}{\arg\max}\, p(z \mid x_k) P(x_k), \qquad (2)$$

respectively. They minimize the sequence and symbol error probability, respectively. In case of communication systems with finite memory, they can be implemented through the Viterbi [5] and the BCJR algorithm [3], respectively.

Without loss of generality, in the following we will often assume that modulation symbols $\{x_k\}$ are independent and uniformly distributed (i.u.d.). As a consequence, all a priori probabilities can be safely discarded from the aforementioned MAP strategies and they become perfectly equivalent to the corresponding ML strategies [6].

DETECTION ON INTERSYMBOL INTERFERENCE CHANNELS

The continuous-time ISI channel may be described, assuming the use of a linear modulation, by means of the following complex baseband equation:

$$y(t) = \sum_k x_k q(t - kT) + w(t), \qquad (3)$$

where $q(t)$ is the received pulse, $w(t)$ is complex white Gaussian noise with two-sided spectral density N_0, and T is the symbol time.

In 1972, Forney showed that ML detection of x can be carried out by an application of the Viterbi algorithm (VA) [2]. Forney first applied a matched filter and sampling operation to the signal $y(t)$ to form the discrete-time model $y_k = \int_{-\infty}^{\infty} y(t) q^*(t - kT)\,dt$. Each sample y_k can be expressed as

$$y_k = \sum_{\ell=-L}^{L} g_\ell x_{k-\ell} + n_k, \qquad (4)$$

where $g_\ell = \int_{-\infty}^{\infty} q(t) q^*(t - \ell T)\,dt$. The variable L specifies the memory of the system and is the smallest value such

that $g_\ell = 0, |\ell| > L$. The noise in the model (4) is not white, but is correlated according to $\mathbb{E}[n_{k+\ell}n_k^*] = N_0 g_\ell$. To obtain white noise, Forney filtered the variables y_k with a whitening filter $\{f_k\}$, which yields

$$r_k = \sum_\ell f_\ell y_{k-\ell} = \sum_{\ell=0}^{L} h_\ell x_{k-\ell} + w_k. \quad (5)$$

In the model (5), the zero-mean noise variables $\{w_k\}$ are uncorrelated with variance N_0. The channel impulse response is causal and is related to $\{g_\ell\}$ as $g_\ell = \sum_{k=0}^{L} h_{k+\ell} h_k^*$, i.e., $g = [g_{-L}, \ldots, g_L]$ is the autocorrelation sequence of $h = [h_0, \ldots, h_L]$. Both samples $\{y_k\}$ and $\{r_k\}$ represent a sufficient statistic and can thus be employed for detection. Throughout this lecture note, the three letters (y, g, n) imply that we are discussing the model (4), while we are discussing (5) if we use (r, h, w).

Forney next observed that each sample r_k only depends on the current channel input x_k and the L most recent channel inputs $x_{k-1}, \ldots, x_{k-L}$. Therefore, the signal can be described by means of a trellis where each state is defined as $\sigma_k = (x_{k-1}, \ldots, x_{k-L})$. Thus, the number of states is $|\chi|^L$. As an example, when $\chi = \{0, 1\}$ and $L = 2$, a section of the corresponding trellis between the discrete-time instants k and $k + 1$ is shown in Figure 1. In this figure, trellis transitions driven by symbol $x_k = 0$ are denoted by using dashed lines, whereas solid lines correspond to transitions driven by $x_k = 1$.

Due to the fact that samples r_k are conditionally independent, the conditional probability density function $p(r|x)$ required for the implementation of the strategy (1) can be expressed in a recursive factorization of the form

$$p(r|x)$$
$$= \prod_k \frac{1}{\pi N_0} \exp\left(-\frac{\left| r_k - \sum_{\ell=0}^{L} h_\ell x_{k-\ell} \right|^2}{N_0} \right). \quad (6)$$

Based on (6) it is straightforward to set up the VA. In fact, under the assumption of i.u.d. modulation symbols and taking into account that the logarithm is a monotonic function, the strategy (1) can be expressed as

$$\hat{x} = \operatorname*{argmax}_x p(r|x)$$
$$= \operatorname*{argmax}_x \ln p(r|x) = \operatorname*{argmin}_x \sum_k \mu_k, \quad (7)$$

where

$$\mu_k = \left| r_k - \sum_{\ell=0}^{L} h_\ell x_{k-\ell} \right|^2 \quad (8)$$

is the so-called branch metric of the VA.

In turbo equalization applications [7], one may resort to the MAP symbol detection strategy. In this case, it is sufficient to replace the VA with the BCJR algorithm, possibly implemented in the logarithmic domain [8]. It will make use of the same branch metric μ_k.

However, a demodulator may just as well take as starting point the model (4) as already shown by Ungerboeck in 1974 in [1]. The model is commonly referred to as the *Ungerboeck model*, while the white-noise model (5) is referred to as the *Forney model*—a nomenclature we will follow. The noise variables $\{n_k\}$ are still Gaussian, but are colored. However, the noise color is irrelevant since the critical issue for the application of a

σ_k
(0, 0)
(1, 0)
(0, 1)
(1, 1)
k $k+1$
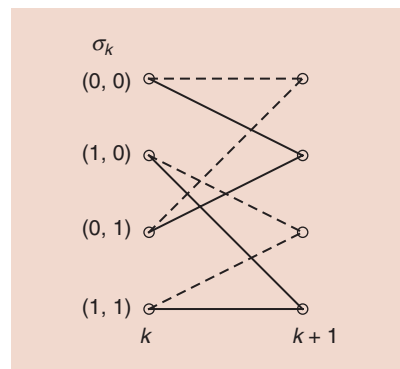
[FIG1] An example of a trellis section.

VA-type-detector is that the conditional pdf $p(r|x)$ has a recursive factorization that can be expressed in terms of the signal y. This is indeed the case as can be seen by expanding (6) (see [1] for further details). See (9) in the box at the bottom of the page, where $\gamma_1 = \prod_k (\pi N_0)^{-1}$, $\gamma_2 = \gamma_1 \exp(-\|r\|^2/N_0)$, and where we used $y_k = \sum_{\ell=0}^{L} h_\ell^* r_{k+\ell}$ in the last equality. Note that γ_2 is independent of $\{x_k\}$ and can be neglected. Again, under the assumption of i.u.d. modulation symbols, the strategy (1) can be expressed as

$$\hat{x} = \operatorname*{argmax}_x \ln p(r|x)$$
$$= \operatorname*{argmax}_x \sum_k \eta_k, \quad (10)$$

where this time the branch metric is

$$\eta_k = \mathcal{R}\left\{ y_k x_k^* - \frac{1}{2}|x_k|^2 g_0 - x_k^* \sum_{\ell=1}^{L} g_\ell x_{k-\ell} \right\}. \quad (11)$$

Again, we see that only the L most recent channel inputs $x_{k-1}, \ldots, x_{k-L}$ are needed at each time epoch k. Ungerboeck's and

$$p(r|x) = \gamma_1 \exp\left(-\frac{\sum_k |r_k|^2}{N_0} + \frac{2}{N_0} \mathcal{R}\left\{ \sum_k \sum_{\ell=0}^{L} x_{k-\ell}^* h_\ell^* r_k \right\} - \frac{1}{N_0} \sum_{k,\ell,\ell'} x_{k-\ell}^* x_{k-\ell'} h_\ell^* h_{\ell'} \right)$$

$$= \gamma_2 \exp\left(\frac{2}{N_0} \mathcal{R}\left\{ \sum_k x_k^* \sum_{\ell=0}^{L} h_\ell^* r_{k+\ell} \right\} - \frac{1}{N_0} \sum_k x_k^* \sum_{\ell=-L}^{L} x_{k-\ell} g_\ell \right)$$

$$= \gamma_2 \prod_k \exp\left(\frac{2}{N_0} \mathcal{R}\left\{ y_k x_k^* - \frac{1}{2}|x_k|^2 g_0 - x_k^* \sum_{\ell=1}^{L} g_\ell x_{k-\ell} \right\} \right) \quad (9)$$

[lecture **NOTES**]

Forney's ML sequence detectors involve different computations, but they traverse the very same trellis, and their final outputs are identical. Two strong aspects of the Ungerboeck model are that no squaring operations are needed and that no whitening is needed.

It is interesting to note that while the BCJR already became available for the Forney model by 1974 with [3], the story differs remarkably for the Ungerboeck model. An equivalent algorithm to the BCJR that operates on the Ungerboeck model and employs the same branch metric (11) was demonstrated as late as 2005 in [4]. As a consequence, turbo equalization based on Ungerboeck's model was not available before 2005. The Ungerboeck model has a number of strengths and has been rediscovered several times during the recent past. Before we turn our attention to three short examples that illuminate its strengths, we first extend it into a model for general linear channels.

DETECTION ON GENERAL LINEAR CHANNELS

Let us write (5) as a matrix equation,

$$r = Hx + w. \qquad (12)$$

In the ISI case, the matrix H is a Toeplitz matrix that represents the convolutional operator. However, (12) can also represent any other linear channel, such as multiple-input, multiple-output (MIMO), intercarrier interference (ICI), MIMO-ISI etc. Irrespective of from where the model (12) came, the corresponding conditional pdf has expression

$$p(r|x) = \gamma_1 \exp\left(-\frac{\|r - Hx\|^2}{N_0}\right). \quad (13)$$

One can reach a tree structure, suitable for demodulation, by a QL factorization $H = QL$ of the channel, which makes the model (12) "causal" in the vector index. This gives

$$\tilde{r} = Q^*r = Lx + \tilde{w},$$

which enables a recursive factorization, similar to (6),

$$p(\tilde{r}|x)$$
$$= \prod_k \frac{1}{\pi N_0} \exp\left(-\frac{\left|\tilde{r}_k - \sum_{\ell=1}^{k} L_{k,\ell}x_\ell\right|^2}{N_0}\right). \quad (14)$$

A tree search procedure can now be reached. In the case of a channel with finite memory, i.e., $L_{k,\ell} = 0, |k - \ell| > L$, the VA or the BCJR can be applied since the tree collapses into a trellis with $|\mathcal{X}|^L$ states. Hence, we refer to (12) as the Forney model for a linear channel.

To avoid computation of a QL factorization of the channel matrix, [9] proposed to first multiply the vector r with a matched filter $y = H^*r = Gx + n$, where $G = H^*H$, and n is colored Gaussian noise with covariance matrix N_0G. However, in view of (4), this is nothing but an extension of the Ungerboeck model for ISI into a formulation for a general linear channel. Next, [9] proceeds with the derivation of a recursive factorization, suitable for a tree search, and finally obtains

$$p(r|x) = \gamma_2 \prod_k \exp\left(\frac{2}{N_0}\mathcal{R}\left\{x_k^*y_k - \frac{1}{2}|x_k|^2 G_{k,k} - x_k^* \sum_{\ell=1}^{k} G_{\ell,k}x_\ell\right\}\right), \qquad (15)$$

where, again, γ_2 is irrelevant for decision. This we recognize as the extension of Ungerboeck's (9) into a formulation for general linear channels.

PROBLEM STATEMENT

Is the choice of model relevant? Is there any example of practical systems where the Ungerboeck model is more convenient? We now give a few examples of systems where the Ungerboeck model can offer superior performance and/or lower complexity.

SOLUTION

CHANNEL SHORTENING DETECTION

Since the VA is often of prohibitive complexity, Falconer and Magee [10] proposed in 1973 to make use of the following reduced-complexity scheme: 1) filter the signal (5) with a filter that aims at reducing the memory of the effective impulse response from L to $K < L$ and 2) apply the VA to the filtered signal, but based on the shorter effective channel. Thus, the VA traverses a trellis with $|\mathcal{X}|^K$ states rather than the full trellis of size $|\mathcal{X}|^L$. In terms

of a general linear channel, what is done is that the conditional pdf (13) is replaced by the mismatched version

$$T(r|x) \propto \exp\left(-\frac{\|Wr - Fx\|^2}{N_0}\right), \quad (16)$$

where W is the channel shortener, F is a matrix that has K nonzero consecutive diagonals, and the normalization constant has been neglected. This specifies a trellis with $|\mathcal{X}|^K$ states so that the VA or the BCJR can be applied. The operations of such VAs or BCJR algorithms are specified by (14) with $\tilde{r}$ and L being replaced by Wr and F, respectively.

However, instead of using (14), we can, with identical complexity, use (15). By expanding the square magnitude in (16) and neglecting the irrelevant terms we can express $T(r|x)$ as

$$T(r|x) \propto \exp\left(\frac{2\mathcal{R}\{x^*F^*Wr\} - x^*F^*Fx}{N_0}\right) \qquad (17)$$

and then execute the trellis processing via (15) by replacing y and G with F^*Wr and F^*F, respectively.

If the processing is done via (15), only the matrices F^*W and F^*F are relevant, and not the matrices W and F themselves. We can therefore relax the structure of F^*W and F^*F so that we replace $T(r|x)$ in (17) with

$$T(r|x) \propto \exp\left(\frac{2\mathcal{R}\{x^*H_r r\} - x^*G_r x}{N_0}\right), \qquad (18)$$

where H_r is arbitrary and G_r is a Hermitian matrix with only the main $2K + 1$ diagonals holding nonzero values. The strength of replacing F^*F with G_r is that the matrix G_r needs not to be positive semidefinite, unlike the matrix F^*F which is positive semidefinite by construction. This allows for a wider class of mismatched conditional pdfs than what can be reached by (16); based on (16), one is restricted to have a positive semidefinite G_r matrix.

As far as the derivation of G_r and H_r according to a proper optimality criterion

is concerned, we refer the reader to [11]. Importantly, to find optimal G_r and H_r is much simpler than finding optimal W and F. Somewhat surprisingly, the optimal G_r matrix to choose is often in-definite so that a mismatched conditional pdf of the form (16) is inferior to the form (18).

In brief, channel shortening has been studied since 1973, but the starting point has always been the Forney model. This is suboptimal, as the optimal solution for an Ungerboeck-based channel shortening receiver can not, in general, be reached with the Forney model.

MAX-LOG-MAP DEMODULATION OF MIMO CHANNELS

The computation of the pdf (13) required for the implementation of the strategies (1) or (2) requires the computation of metrics $\|r - Hx\|^2$ for all possible values of x. How many complex multiplications are needed to do this task? If we assume an $M \times M$ channel matrix, we have $|\mathcal{X}|^M$ vectors x to test. For each vector we need M^2 multiplications to form Hx, and then M more to compute the norm. Hence, a brute force evaluation would give about $|\mathcal{X}|^M (M+1)M$ complex multiplications. In [12], a much more efficient computation is presented by a clever rewriting of the associated terms of computing each metric $\|r - Hx\|^2$. To exemplify how the metric is rewritten for simplifying the calculations, we consider the case of a 2×2 MIMO system. The received signal, the channel matrix, and the transmitted data vector, all complex valued, are expressed as

$$r = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \quad H = \begin{bmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{bmatrix}$$
$$x = \begin{bmatrix} \mathcal{R}\{x_1\} + j\mathcal{I}\{x_1\} \\ \mathcal{R}\{x_2\} + j\mathcal{I}\{x_2\} \end{bmatrix},$$

respectively. With that, in [12] the metric is expressed in the following manner [see (19) in the box at the bottom of the page].

Notice that we neglected the term $\|r\|^2$, which is irrelevant for detection. Although never mentioned in [12], this is precisely the Ungerboeck model, but in a real-valued formulation. In fact, as defined in the section "Detection on General Linear Channels," it is $y = H^* r$ and $G = H^* H$, and thus $y_1^* = r_1^* H_{1,1} + r_2^* H_{2,1}$ and $y_2^* = r_1^* H_{1,2} + r_2^* H_{2,2}$, $G_{1,1} = |H_{1,1}|^2 + |H_{2,1}|^2$, $G_{1,2} = H_{1,1}^* H_{1,2} + H_{2,1}^* H_{2,2} = G_{2,1}^*$, etc. Hence, (19) becomes

$$\|r - Hx\|^2 \propto \begin{pmatrix} -2\mathcal{R}\{y_1^*\} \\ 2\mathcal{I}\{y_1^*\} \\ -2\mathcal{R}\{y_2^*\} \\ 2\mathcal{I}\{y_2^*\} \end{pmatrix}^T \begin{pmatrix} \mathcal{R}\{x_1\} \\ \mathcal{I}\{x_1\} \\ \mathcal{R}\{x_2\} \\ \mathcal{I}\{x_2\} \end{pmatrix} + \begin{pmatrix} G_{1,1} \\ G_{1,1} \\ G_{2,2} \\ G_{2,2} \end{pmatrix}^T \begin{pmatrix} \mathcal{R}\{x_1\}^2 \\ \mathcal{I}\{x_1\}^2 \\ \mathcal{R}\{x_2\}^2 \\ \mathcal{I}\{x_2\}^2 \end{pmatrix} + \begin{pmatrix} 2\mathcal{R}\{G_{1,2}\} \\ -2\mathcal{I}\{G_{1,2}\} \\ 2\mathcal{I}\{G_{1,2}\} \\ 2\mathcal{R}\{G_{1,2}\} \end{pmatrix}^T \begin{pmatrix} \mathcal{R}\{x_1\}\mathcal{R}\{x_2\} \\ \mathcal{R}\{x_1\}\mathcal{I}\{x_2\} \\ \mathcal{I}\{x_1\}\mathcal{R}\{x_2\} \\ \mathcal{I}\{x_1\}\mathcal{I}\{x_2\} \end{pmatrix},$$

(20)

which is exactly the argument of the exponential in (15) in real-valued formulation.

Based on the formulation (19), [12] developed a methodolgy to calculate all the $|\mathcal{X}|^M$ metrics by only $2M^2(2M+3) - 2M$ multiplications. This remarkable result relies on the structure of the Ungerboeck model (15), which allows for a hierarchical formulation of the minimum metric terms, using submetrics in an efficient manner avoiding duplication of calculations. The Ungerboeck metric is utilized in such a way that parallelization is achieved and multiple calculations are done in one clock cycle. Furthermore, a doubly recursive evaluation of submetrics is used; for a detailed description of this, see [12]. In [12], it is also shown that an efficient implementation of a soft output max-log MAP detector for a $2 \times M$ MIMO system with quadrature amplitude modulation (QAM) inputs reduces the number of candidate tests by a factor of $|\mathcal{X}|$ by rewriting the minimum metric expression in a hierarchical manner. The remaining candidate tests are performed in a recursive fashion avoiding multiplications altogether. As a result, the computational complexity for the metric calculations has been reduced by a factor of 250 for a $2 \times M$ MIMO system with Gray-coded 64-ary QAM (64QAM). Furthermore, it was estimated that with 10-bit quantization of the metric component values, 64QAM, and a 2×2 MIMO system, a chip area of 0.031 mm^2 would be required for a clock frequency of 125 MHz and 65-nm complementary metal–oxide–semiconductor (CMOS) technology. For more details and applications to the IEEE 802.11n standard, see [12].

FACTOR-GRAPH-BASED DETECTOR WITH LINEAR COMPLEXITY IN THE NUMBER OF INTERFERERS

As said earlier, optimal detection by means of the VA or the BCJR algorithm works over a trellis with $|\mathcal{X}|^L$ states. Channel shortening tries to transform the original channel into a shorter one before detection. In addition to channel shortening, other approaches to complexity reduction have been based on a reduced search over the original trellis or on a search over a reduced trellis obtained from the original one through a partial representation of the symbols in the trellis state definition. Many

$$\|r - Hx\|^2 = (r - Hx)^*(r - Hx) \propto -2\mathcal{R}\{r^* Hx\} + x^* H^* Hx$$
$$= \begin{pmatrix} -2\mathcal{R}\{r_1^* H_{1,1} + r_2^* H_{2,1}\} \\ 2\mathcal{I}\{r_1^* H_{1,1} + r_2^* H_{2,1}\} \\ -2\mathcal{R}\{r_1^* H_{1,2} + r_2^* H_{2,2}\} \\ 2\mathcal{I}\{r_1^* H_{1,2} + r_2^* H_{2,2}\} \end{pmatrix}^T \begin{pmatrix} \mathcal{R}\{x_1\} \\ \mathcal{I}\{x_1\} \\ \mathcal{R}\{x_2\} \\ \mathcal{I}\{x_2\} \end{pmatrix} + \begin{pmatrix} |H_{1,1}|^2 + |H_{2,1}|^2 \\ |H_{1,1}|^2 + |H_{2,1}|^2 \\ |H_{1,2}|^2 + |H_{2,2}|^2 \\ |H_{1,2}|^2 + |H_{2,2}|^2 \end{pmatrix}^T \begin{pmatrix} \mathcal{R}\{x_1\}^2 \\ \mathcal{I}\{x_1\}^2 \\ \mathcal{R}\{x_2\}^2 \\ \mathcal{I}\{x_2\}^2 \end{pmatrix}$$
$$+ \begin{pmatrix} 2\mathcal{R}\{H_{1,1}^* H_{1,2} + H_{2,1}^* H_{2,2}\} \\ -2\mathcal{I}\{H_{1,1}^* H_{1,2} + H_{2,1}^* H_{2,2}\} \\ 2\mathcal{I}\{H_{1,1}^* H_{1,2} + H_{2,1}^* H_{2,2}\} \\ 2\mathcal{R}\{H_{1,1}^* H_{1,2} + H_{2,1}^* H_{2,2}\} \end{pmatrix}^T \begin{pmatrix} \mathcal{R}\{x_1\}\mathcal{R}\{x_2\} \\ \mathcal{R}\{x_1\}\mathcal{I}\{x_2\} \\ \mathcal{I}\{x_1\}\mathcal{R}\{x_2\} \\ \mathcal{I}\{x_1\}\mathcal{I}\{x_2\} \end{pmatrix}. \quad (19)$$

[lecture **NOTES**]
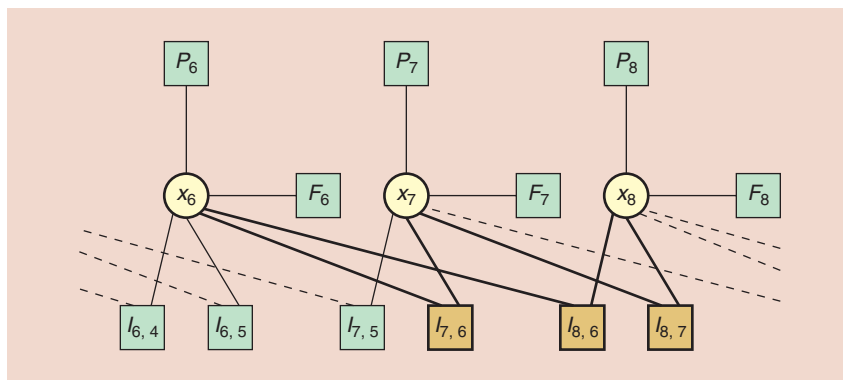
[FIG2] Three sections of the FG corresponding to (25), for the case $L = 2$.

papers have investigated these approaches (as an example, see [13] and the references therein). They all have in common that they work on the Forney observation model. An attempt with scarce success has been tried in [14] to adapt some of them to the Ungerboeck observation model. The reason is related to the fact that the partial metric of the VA algorithm does not have, in the case of the Ungerboeck observation model, a probabilistic meaning.

However, the Ungerboeck model allows a different approach to complexity reduction [15]. Neglecting the factors in (9) that are irrelevant for detection, we can write

$$p(r \mid x) \propto \prod_k \exp\left(\frac{2}{N_0}\mathcal{R}\left\{y_k x_k^*\right.\right.$$
$$\left.\left. -\frac{1}{2}\mid x_k \mid^2 g_0 - x_k^* \sum_{\ell=1}^{L} g_\ell x_{k-\ell}\right\}\right)$$
$$(21)$$

$$= \prod_k F_k(x_k) \prod_{\ell=1}^{L} I_{k,k-\ell}(x_k, x_{k-\ell})$$
$$(22)$$

having defined

$$F_k(x_k)$$
$$= \exp\left(\frac{2}{N_0}\mathcal{R}\left\{y_k x_k^* - \frac{1}{2}\mid x_k \mid^2 g_0\right\}\right) \quad (23)$$

$$I_{k,k-\ell}(x_k, x_{k-\ell})$$
$$= \exp\left(-\frac{2}{N_0}\mathcal{R}\left\{g_\ell x_k^* x_{k-\ell}\right\}\right). \quad (24)$$

Thus, the joint a posteriori probability of the transmitted symbols can be factorized as

$$P(x \mid r) \propto \prod_k P_k(x_k) F_k(x_k)$$
$$\prod_{\ell=1}^{L} I_{k,k-\ell}(x_k, x_{k-\ell}), \quad (25)$$

where $P_k(x_k)$ is the a priori probability of symbol x_k.

The factorization (25) can be visualized through an FG; an example is given in Figure 2. In this graph, variable and factor nodes are represented through circles and squares, respectively. An edge connects a variable node x_k with a factor node if and only if that variable is an argument of the factor corresponding to that factor node. In the figure, we used dashed lines to represent edges involving nodes not explicitly represented in the graph. The meaning of bold edges will be explained below. Note that, when $g_\ell = 0$, the factor $I_{k,k-\ell}$ is equal to one and can thus be dropped from the factorization (25). In practice, the node $I_{k,k-\ell}$ must be included in (25) only when $g_\ell \neq 0$, i.e., only when x_k and $x_{k-\ell}$ interfere with each other.

The factorization (25) is exact, since no approximation was adopted in its derivation. On the other hand, the marginalization of (25), required for computing the a posteriori probabilities $\{P(x_k \mid r)\}$, cannot be exactly carried out by applying the SPA to the FG in Figure 2, since it contains cycles. One of these cycles is indicated in the figure in bold. It is easy to prove that the FG corresponding to (25) cannot contain any cycle of length lower than six, irrespective of the number of symbols that interfere with each other. In fact, being factor nodes of at most degree two, the necessary and sufficient condition for the arising of a cycle of length four is to have two factor nodes of degree two connected to the same couple of variable nodes, and this is clearly not possible, by definition of $I_{k,k-\ell}$. Hence, in this

case, the SPA may lead to favorable results since it is generally expected to provide a good approximation of the exact marginalizations when the length of the cycles is at least six.

The algorithm resulting from the application of the SPA to the described FG is iterative and has a complexity per iteration, which is linear in the number of interferers. This is related to the adopted factorization having the appealing property that nodes $I_{k,k-\ell}(x_k,x_{k-\ell})$, whose number linearly increases with the number of interferers, have degree two (i.e., they have two edges) irrespective of the number of interferers.

CONCLUSIONS

Although the Ungerboeck and the Forney observation models are equivalent whenever optimal ML receivers are employed, the two models have different properties with suboptimal receivers. Almost all reduced-complexity receivers take the Forney model as the basis for complexity reduction. The best example is that it took more than three decades from the time that the Ungerboeck model was published until a BCJR was derived for it. Thus, no reduced-complexity Ungerboeck-based BCJRs could have been researched until only recently. Meanwhile, the amount of research devoted to reduced-complexity Forney-based BCJRs is impressive. We believe that many algorithms can benefit from being implemented in the Ungerboeck model and that there is much to gain if the awareness of the model is increased. As a step in this direction, we have discussed three examples where the key building block is the Ungerboeck model.

AUTHORS
Fredrik Rusek (fredrik.rusek@eit.lth.se) is an associate professor in the Department of Electrical and Information Technology at Lund University, Sweden.

Giulio Colavolpe (giulio.colavolpe@unipr.it) is an associate professor in the Dipartimento di Ingegneria dell'Informazione at the University of Parma, Italy.

Carl-Erik W. Sundberg (cews@ieee.org) is the president and chief scientist at SundComm, Sunnyvale, California. From 1981 to 2000, he was a distinguished

member of technical staff at Bell Laboratories, Murray Hill, New Jersey.

REFERENCES

[1] G. Ungerboeck, "Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems," *IEEE Trans. Commun.*, vol. 22, no. 5, pp. 624–636, May 1974.

[2] G. D. Forney, Jr., "Maximum likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 18, no. 3, pp. 363–378, May 1972.

[3] L. R. Bahl, J. Cocke, F. Jelinek, and R. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, Mar. 1974.

[4] G. Colavolpe and A. Barbieri, "On MAP symbol detection for ISI channels using the Ungerboeck observation model," *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 720–722, Aug. 2005.

[5] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.

[6] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York: McGraw-Hill, 2008.

[7] C. Douillard, M. Jézequel, and C. Berrou, "Iterative correction of intersymbol-interference: Turbo-equalization," *Eur. Trans. Telecommun.*, vol. 6, no. 5, pp. 507–511, Sept. 1995.

[8] P. Roberston, E. Villebrun, and P. Hoeher, "Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding," *Eur. Trans. Telecommun.*, vol. 8, no. 2, pp. 119–125, Mar./Apr. 1997.

[9] C. Kuhn and N. Goertz, "A low-complexity path metric for tree-based multiple-antenna detectors," in *Proc. IEEE Int. Conf. Communications*, May 2007, pp. 1024–1029.

[10] D. D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell Syst. Tech. J.*, vol. 52, no. 9, pp. 1541–1562, Nov. 1973.

[11] F. Rusek and A. Prlja, "Optimal channel shortening of MIMO and ISI channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 810–818, Feb. 2012.

[12] N. Graef, J. Hammerschmidt, and C.-E. W. Sundberg, "A low-complexity max-log-MAP detector," *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2251–2254, Aug. 2009.

[13] A. Prlja and J. B. Anderson, "Reduced-complexity receivers for strongly narrowband intersymbol interference introduced by faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2591–2601, Sept. 2012.

[14] A. Hafeez and W. E. Stark, "Decision feedback sequence estimation for unwhitened ISI channels with applications to multiuser detection," *IEEE J. Select. Areas Commun.*, vol. 16, no. 9, pp. 1785–1795, Dec. 1998.

[15] G. Colavolpe, D. Fertonani, and A. Piemontese, "SISO detection over linear channels with linear complexity in the number of interferers," *IEEE J. Select. Topics Signal Processing,* vol. 5, pp. 1475–1485, Dec. 2011.

[SP]

[special **REPORTS**] (continued from page 15)

consideration. The separation of sea level and land motion change is a matter of great importance for global change research, Löfgren says. "How much does the sea level change in different parts of the world and what are the causes of this change?"

The researchers note that existing coastal GNSS stations, installed primarily for the purpose of measuring land movements, can be easily adapted to make sea level measurements. "We have successfully tested a method where only one of the antennas is used to receive the radio signals," Löfgren says. "That means that existing coastal GNSS stations—there are hundreds of them all over the world—can also be used to measure the sea level."

Löfgren regards signal processing as essential to his research. "What I want to do is to convert my GNSS measurements into measurements of sea level in the most accurate way possible," he remarks. "Most of the signal processing is more or less standard in the GNSS world, but I have applied it on a new and different data set."

For the two-antenna technique, Löfgren determines the vertical distance between the upward-looking and the downward-looking antenna (the downward-looking antenna will appear to be a virtual antenna below the sea level, since the reflected signal will travel an additional path compared to the direct signal). "The signal processing is done by analysis of the phase of the recorded signals," he says. "An observational model is set up for the difference in recorded phase between the two antennas (incorporating clock differences in the receivers, differences in geometry and differences in the phase ambiguity parameter), and it is then fitted in a least squares sense to the phase observations."

For the one-antenna technique, a different type of signal processing is applied. "The interference between the direct and the reflected signals can be seen as oscillations in the signal-to-noise ratio (SNR) observable," Löfgren says. "With the assumption of a horizontal non-moving sea level, the frequency of these oscillations is constant with respect to the sine of the satellite elevation angle." This means that the oscillations first need to be found and extracted from the data. Next, the oscillations' frequency content (with respect to the sine of the satellite elevation angle instead of the usual time) should be found either by Fourier transform or a Lomb-Scargle periodogram (LSP), Löfgren says. Finally, the main oscillation frequency must be converted to the distance between the antenna and the reflection point, which is directly proportional to the sea level.

"In both the one- and two-antenna methods, the actual installations that measures reflected signals are already set up," Löfgren says. This means that the geodetic GNSS receivers are first applying some kind of signal processing when they record the satellite signals. "What I am using as techniques are least squares analysis [for the] two-antenna technique, and LSP [for the] one-antenna technique."

For the project's next step, the researchers are looking toward developing multi-GNSS solutions, possibly even combining GPS and GLONASS signals together to increase the number of observations in a combined phase delay analysis, providing more accurate sea level estimates. The combination of GPS and GLONASS for SNR analysis is expected to increase the temporal resolution of the corresponding sea level results.

After that step is accomplished, the goal will be to use multi-GNSS, multi-frequency, phase delay, and SNR analysis in a filter approach. "Doing so, we expect that it will be possible to derive continuous and accurate absolute GNSS sea level time series in a wide range of wind speeds," Löfgren says.

AUTHOR

John Edwards (jedwards@johnedwardsmedia.com) is a technology writer based in the Phoenix, Arizona, area.

[SP]

sp **TIPS&TRICKS**

Rodrigo Capobianco Guido

Practical and Useful Tips on Discrete Wavelet Transforms

The discrete wavelet transform (DWT) [1] is one of the most powerful tools for time-frequency signal analysis. Its applicability is extremely relevant in various areas of science, as exemplified in [2], with digital signal processing (DSP) as the most notable one. After teaching this topic for many years, I have noted that neither young DSP students nor experienced researchers have perceived several interesting aspects of the DWT from a practical point of view. Thus, the objective of this article is to construe such relevant aspects, providing useful tips to calculate the transform in one (1-D) and two (2-D) dimensions. The entire discussion is also valid to the discrete wavelet-packet transform (DWPT), which extends the decomposition carried out by the DWT so that a finer time-frequency analysis takes place, and also to the discrete shapelet transform (DST) [3], which extends the properties of the DWT and DWPT so that a joint time-frequency-shape signal analysis becomes possible.

THE TIPS

The tips given herein come after a short review, which is necessary to understand them.

1-D DWT CALCULATION

Mallat's algorithm [4] is the commonly used method to calculate the DWT of a certain discrete-time signal ($f_{[\cdot]}$) of length N. The procedure, consisting of a simple matrix multiplication, spans both the required convolutions of $f_{[\cdot]}$ with $h_{[\cdot]}$ and with $g_{[\cdot]}$, which represent, respectively, a low-pass filter and a high-pass filter, and the downsamplings by two, all possible due to the reduction of the signals bandwidths

[1]. Filters $h_{[\cdot]}$ and $g_{[\cdot]}$ form the analysis filter pair, which have, most of the time, the same length M, and form a quadrature mirror filter (QMF) [4] set that presents an orthogonality condition and half-band cutoff frequencies. The complete process to transform $f_{[\cdot]}$ is shown in the boxed equation at the bottom of the page.

As the matrix $A_{[\cdot][\cdot]}$, formed by the filters coefficients, advances from the first pair of lines until the last one, a shift to the right becomes necessary so that $h_{[\cdot]}$ and $g_{[\cdot]}$ start to be written two positions ahead in each subsequent pair. In case some of the coefficients fall beyond the length of a row, e.g., at the bottom of matrix $A_{[\cdot][\cdot]}$, they are pushed back at the beginning of the same row and the remaining positions are set to zero. This is known as *wraparound*. The procedure explained consists of the way the convolutions and the downsamplings were implemented.

After performing the calculations above, the resulting DWT corresponds to the concatenation of the subsignal formed by the even indexes of $r_{[\cdot]}$, designated *approximation* of length ($N/2$), followed by that formed by the odd indexes of $r_{[\cdot]}$, designated *detail* of length ($N/2$). Thus, DWT $(f_{[\cdot]}) = \{r_0, r_2, r_4, r_{N-2}..., r_1,$

$r_3, r_5, ..., r_{N-1}\}$, registering the same length of the input signal, i.e., N. This corresponds to the first-level DWT ($f_{[\cdot]}$). If the detail subsignal is kept intact and the approximation subsignal is considered as being a new input to the same algorithm, then two other subsignals of length ($N/4$) are obtained. Their concatenation, following the same process described above, generates one subsignal of length ($N/2$) that replaces the original input of the same length. The new complete signal is of length N and corresponds to the second-level DWT ($f_{[\cdot]}$). The process can be repeated ($\log(N)/\log(2)$) times, i.e., until the length of the approximation subsignal at the current level becomes 1.

The aforementioned process is one we can easily find in literature, however, if the reader tries to calculate a DWT by hand, to check if he or she has learned the algorithm correctly, or even try to implement a computer software to perform the calculations, a problem may appear: as the level of decomposition advances, the approximation subsignal, used as input, reduces its length to half. How can one proceed with the calculations if the dimension of the input becomes lesser than the filters support-size? For example, letting $N = 8$ and $M = 4$, then, after the second-level

$$
\underbrace{\begin{pmatrix}
h_0 & h_1 & h_2 & \cdots & \cdots & \cdots & \cdots & \cdots & h_{M-1} & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
g_0 & g_1 & g_2 & \cdots & \cdots & \cdots & \cdots & \cdots & g_{M-1} & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
0 & 0 & h_0 & h_1 & h_2 & \cdots & \cdots & \cdots & \cdots & h_{M-1} & 0 & \cdots & \cdots & \cdots & 0 \\
0 & 0 & g_0 & g_1 & g_2 & \cdots & \cdots & \cdots & \cdots & g_{M-1} & 0 & \cdots & \cdots & \cdots & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
h_2 & h_3 & \cdots & \cdots & \cdots & \cdots & h_{M-1} & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & h_0 & h_1 \\
g_2 & g_3 & \cdots & \cdots & \cdots & \cdots & g_{M-1} & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & g_0 & g_1
\end{pmatrix}}_{\text{matrix}\, A_{[\cdot][\cdot]}}
\cdot
\underbrace{\begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \\ \cdots \\ \cdots \\ \cdot \\ \cdot \\ \cdot \\ f_{N-2} \\ f_{N-1} \end{pmatrix}}_{\text{input}(f_{[\cdot]})}
=
\underbrace{\begin{pmatrix} r_0 \\ r_1 \\ r_2 \\ \cdots \\ \cdots \\ \cdots \\ \cdot \\ \cdot \\ \cdot \\ r_{N-2} \\ r_{N-1} \end{pmatrix}}_{\text{output}(r_{[\cdot]})}
\cdot
$$

Signal and Information Processing: The Heartbeat of a Smart Society

IEEE
IEEE Signal Processing Society

Shanghai, China
ICASSP·2016
41st IEEE International Conference on Acoustics, Speech and Signal Processing, 20-25 March 2016

General Chairs
Zhi Ding, Univ. of California, Davis, USA
Zhi-Quan Luo, Univ. of Minnesota, USA
Wenjun Zhang, Shanghai Jiao Tong Univ., China

Technical Program Chairs
P. C. Ching, Chinese Univ. of Hong Kong, HK
Dominic K.C. Ho, Univ. of Missouri, USA

Finance Chairs
Shuguang Cui, Texas A&M Univ., USA
Rong Xie, Shanghai Jiao Tong Univ., China

Plenaries Chairs
Zhi-Pei Liang, UIUC, USA
Björn Ottersten, Univ. of Luxembourg, Luxembourg

Special Sessions Chairs
Tim Davidson, McMaster Univ., Canada
Jianguo Huang, Northwestern Polytech. Univ., China

Tutorials Chairs
Jian Li, Univ. of Florida, USA
Jose Principe, Univ. of Florida, USA

Student Session Chair
Wei Zhang, Univ. of New South Wales, AU

Registration Chairs
Tongtong Li, Michigan State Univ., USA
Xiaojun Yuan, ShanghaiTech Univ., China

Publicity Chairs
Xiaokang Yang, Shanghai Jiao Tong Univ., China
Mounir Ghogho, Leeds Univ., UK
Ignacio Santamaria, Univ. of Cantabria, Spain

Publication Chairs
Min Dong, Univ. of Ontario Inst. of Tech., Canada
Thomas Fang Zheng, Tsinghua Univ., China

Industrial & Exhibit Chairs
Li Deng, Microsoft, USA
Jinyu Li, Microsoft, USA
Cathy Wicks, Texas Instruments, USA

Local Arrangement Chairs
Ning Liu, Shanghai Jiao Tong Univ., China
Meixia Tao, Shanghai Jiao Tong Univ., China

Webmaster
Yi Xu, Shanghai Jiao Tong Univ., China

Workshop Chairs
Jianguo Huang, Northwestern Polytech. Univ., China
Jiwu Huang, Sun Yat-sen Univ., China

ICASSP2016: Signal and information processing is the driving heartbeat in the development of technologies that enrich our lives and advance our society. The 41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP) will be held in the Shanghai International Convention Center, Shanghai, China between March 20 and 25, 2016. The conference provides, both for researchers and developers, an engaging forum to exchange ideas and propel new developments in this field. The 2016 conference will showcase world-class presentations by internationally renowned speakers and will facilitate a fantastic opportunity to network with like-minded professionals from around the world. Topics include but are not limited to:

Audio and acoustic signal processing	Multimedia signal processing
Bio-imaging and biomedical signal processing	Sensor array & multichannel signal processing
Signal processing education	Design & implementation of signal processing systems
Speech processing	Signal processing for communications & networking
Industry technology tracks	Image, video & multidimensional signal processing
Information forensics and security	Signal processing theory & methods
Machine learning for signal processing	Spoken language processing
Signal processing for Big Data	Signal processing for the Internet of Things

Shanghai: Shanghai is the most populous city in China and one of the most populous cities in the world. A global city, Shanghai exerts influence over global commerce, finance, culture, art, fashion, research and entertainment. The city is located in the middle portion of the Chinese coast, and sits at the mouth of the Yangtze River. The city is a tourist destination renowned for its historical landmarks, such as the Bund and City God Temple, and its modern and ever-expanding Pudong skyline including the Oriental Pearl Tower. Today, Shanghai is the largest center of commerce and finance in mainland China, and has been described as the "showpiece" of the world's fastest-growing major economy.

Submission of Papers: Prospective authors are invited to submit full-length papers, with up to four pages for technical content including figures and possible references, and with one additional optional 5th page containing only references. A selection of best student papers will be made by the ICASSP 2016 committee upon recommendations from the Technical Committees.

Tutorial and Special Session Proposals: Tutorials will be held on March 20 and 21, 2016. Tutorial proposals must include title, outline, contact information, biography and selected publications for the presenter(s), and a description of the tutorial and the material to be distributed to participants. Special session proposals must include a topical title, rationale, session outline, contact information, and a list of invited speakers. Additional information can be found at the ICASSP 2016 website.

Signal Processing Letters: Authors of IEEE Signal Processing Letters (SPL) papers will be given the opportunity to present their work at ICASSP 2016, subject to space availability and approval by the ICASSP Technical Program Chairs. SPL papers published between January 1, 2015 and December 31, 2015 are eligible for presentation at ICASSP 2016.

Show and Tell: S&T offers a perfect stage to showcase innovative ideas in all technical areas of interest at ICASSP. S&T sessions contain demos that are highly interactive and visible. Please refer to the ICASSP 2016 website for additional information regarding demo submission.

Important Deadlines:

Special session & tutorial proposals	August 3, 2015
Notification of special session & tutorial acceptance	September 11, 2015
Submission of regular papers	September 25, 2015
Signal processing letters	December 16, 2015
Notification of paper acceptance	December 21, 2015
Revised paper upload	January 22, 2016
Author registration	January 22, 2016

[sp **TIPS&TRICKS**] continued

decomposition, the input subsignal to be used in the third level has length two, i.e, the multiplication is not possible because matrix $A_{[\cdot][\cdot]}$ and the input are not compatible in terms of length. Some of us may rapidly respond: proceed with a zero-padding in the input so that it reaches, at least, the filters support-size. This solution is, however, equivalent to simply discarding some of the filters coefficients, because they will be multiplied by zero, therefore, it is incorrect. The right solution to this problem is to repeat the input so many times as necessary until it reaches the minimum required length so that the matrix multiplication becomes possible.

NUMERICAL EXAMPLE:
Let $f_{[\cdot]} = \{1, 2, 3, 4, 4, 3, 2, 1\}$, which is a discrete-time signal of length $N = 8$, and let $h_{[\cdot]} = \{(1+\sqrt{3})/(4\sqrt{2}), (3+\sqrt{3})/(4\sqrt{2}), (3-\sqrt{3})/(4\sqrt{2}), (1-\sqrt{3})/(4\sqrt{2})\}$ and $g_{[\cdot]} = \{(1-\sqrt{3})/(4\sqrt{2}), (-3+\sqrt{3})/(4\sqrt{2}), (3+\sqrt{3})/(4\sqrt{2}), (-1-\sqrt{3})/(4\sqrt{2})\}$, which correspond to the Daubechies-4 filter pair [4]. The first and second levels of decomposition produce, respectively, the results shown in the boxed equation at the top of the page.

Performing the third-level decomposition, by using the subsignal $\{(156 + 28\sqrt{3})/32, (164 - 28\sqrt{3})/32\}$ as input, yields the boxed equation at the bottom of the page, which corresponds to the correct procedure. The reader can note that, when this tip is applied, $A_{[\cdot][\cdot]}$ is not a squared matrix. Instead, it contains only the required number of lines to obtain the number of expected coefficients in the resulting signal, i.e., two in this example. The calculations above result in:

$$DWT_{level3} = \left\{ 5\sqrt{2}, \frac{7\sqrt{3}-1}{4\sqrt{2}}, \frac{4+28\sqrt{3}}{32}, \frac{-4-28\sqrt{3}}{32}, 0, \frac{2\sqrt{3}}{4\sqrt{2}}, 0, \frac{-2\sqrt{3}}{4\sqrt{2}} \right\}.$$

In case there is a need to invert the transformations at each level, the synthesis filter bank is adopted. It is composed of the filters $\bar{h}_{[\cdot]}$ and $\bar{g}_{[\cdot]}$, which are respectively defined as being $\bar{h}_k = h_{M-k-1}$ and $\bar{g}_k = (-1)^{k+1}h_k$. An easier way to implement the inverse DWT (IDWT) is to note that the

$$DWT_{level1} = \left\{ \frac{20-4\sqrt{3}}{4\sqrt{2}}, \frac{30}{4\sqrt{2}}, \frac{20+4\sqrt{3}}{4\sqrt{2}}, \frac{10}{4\sqrt{2}}, 0, \frac{2\sqrt{3}}{4\sqrt{2}}, 0, \frac{-2\sqrt{3}}{4\sqrt{2}} \right\} \text{ and}$$

$$DWT_{level2} = \left\{ \frac{156+28\sqrt{3}}{32}, \frac{164-28\sqrt{3}}{32}, \frac{4+28\sqrt{3}}{32}, \frac{-4-28\sqrt{3}}{32}, 0, \frac{2\sqrt{3}}{4\sqrt{2}}, 0, \frac{-2\sqrt{3}}{4\sqrt{2}} \right\}.$$

matrix $A_{[\cdot][\cdot]}$ is orthogonal. Therefore, its inverse, which allows the process to be inverted, corresponds exactly to its transpose. Examples can be found in [4].

2-D-DWT CALCULATION
The definition of the 2-D-DWT can be found in [4]. According to it, an N by N matrix, usually representing a digital image, is used as input while the transformation output consists of another matrix with the same dimensions. This output is composed of the concatenation of four submatrices: $a_{[\cdot][\cdot]}$, $p_{[\cdot][\cdot]}$, $v_{[\cdot][\cdot]}$, and $d_{[\cdot][\cdot]}$, which correspond, respectively, to the approximations of the lines followed by the approximations of the columns, to the approximations of the lines followed by the details of the columns, to the details of the lines followed by the approximations of the columns, and to the details of the lines followed by the details of the columns. Assuming that $S_{[\cdot][\cdot]}$ is the $N \times N$ input matrix, its 2-D-DWT is

$$\begin{pmatrix} a_{[\cdot][\cdot]} & p_{[\cdot][\cdot]} \\ \left(\frac{N}{2} \times \frac{N}{2}\right) & \left(\frac{N}{2} \times \frac{N}{2}\right) \\ v_{[\cdot][\cdot]} & d_{[\cdot][\cdot]} \\ \left(\frac{N}{2} \times \frac{N}{2}\right) & \left(\frac{N}{2} \times \frac{N}{2}\right) \end{pmatrix}.$$

An interesting algorithm to calculate 2-D-DWT($S_{[\cdot][\cdot]}$) directly, instead of calculating the four submatrices separately, is the following:
1) obtain $R_{[\cdot][\cdot]} = A_{[\cdot][\cdot]}S_{[\cdot][\cdot]}A_{[\cdot][\cdot]}^{T}$, being $A_{[\cdot][\cdot]}^{T}$ the transpose of $A_{[\cdot][\cdot]}$
2) rearrange the elements of $R_{[\cdot][\cdot]}$ so that
■ $a_{[\cdot][\cdot]}$ is formed by the elements of $R_{[\cdot][\cdot]}$, which are on the even rows and even columns (starting at line and column zero)
■ $p_{[\cdot][\cdot]}$ is formed by the elements of $R_{[\cdot][\cdot]}$, which are on the even rows and odd columns
■ $v_{[\cdot][\cdot]}$ is formed by the elements of $R_{[\cdot][\cdot]}$, which are on the odd rows and even columns
■ $d_{[\cdot][\cdot]}$ is formed by the elements of $R_{[\cdot][\cdot]}$, which are on the odd rows and odd columns.

The rearranged matrix corresponds to 2-D-DWT ($S_{[\cdot][\cdot]}$). Particularly, the aforementioned procedure corresponds to the first-level 2-D-DWT. To obtain the next levels, the submatrix $a_{[\cdot][\cdot]}$ of the current level is used as the new input and the entire process is repeated. In the calculations, matrix $A_{[\cdot][\cdot]}$ is exactly the same defined for the 1-D-DWT. As previously explained, wraparounds and repetitions of the input signal may be required.

The 2-D-IDWT can be obtained by the inverse algorithm, i.e.,
1) rearrange the elements of 2-D-DWT ($S_{[\cdot][\cdot]}$) so that:
■ $a_{[\cdot][\cdot]}$ is now formed by the elements of 2-D-DWT($S_{[\cdot][\cdot]}$), which are on the even rows and even columns
■ $p_{[\cdot][\cdot]}$ is now formed by the elements of 2-D-DWT($S_{[\cdot][\cdot]}$), which are on the even rows and odd columns
■ $v_{[\cdot][\cdot]}$ is now formed by the elements of 2-D-DWT($S_{[\cdot][\cdot]}$), which are on the odd rows and even columns

$$\underbrace{\begin{pmatrix} \frac{1+\sqrt{3}}{4\sqrt{2}} & \frac{3+\sqrt{3}}{4\sqrt{2}} & \frac{3-\sqrt{3}}{4\sqrt{2}} & \frac{1-\sqrt{3}}{4\sqrt{2}} \\ \frac{1-\sqrt{3}}{4\sqrt{2}} & \frac{-3+\sqrt{3}}{4\sqrt{2}} & \frac{3+\sqrt{3}}{4\sqrt{2}} & \frac{-1-\sqrt{3}}{4\sqrt{2}} \end{pmatrix}}_{\text{matrix } A_{[\cdot][\cdot]}} \cdot \begin{pmatrix} \frac{156+28\sqrt{3}}{32} \\ \frac{164+28\sqrt{3}}{32} \\ \frac{156+28\sqrt{3}}{32} \\ \frac{164+28\sqrt{3}}{32} \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\b\end{matrix}}\right\}\text{input} \\ \left.\vphantom{\begin{matrix}a\\b\end{matrix}}\right\}\text{input repeated once} \end{matrix}$$

IEEE TRANSACTIONS ON

SIGNAL AND INFORMATION PROCESSING OVER NETWORKS

Now accepting paper submissions

The new *IEEE Transactions on Signal and Information Processing over Networks* publishes high-quality papers that extend the classical notions of processing of signals defined over vector spaces (e.g. time and space) to processing of signals and information (data) defined over networks, potentially dynamically varying. In signal processing over networks, the topology of the network may define structural relationships in the data, or may constrain processing of the data. Topics of interest include, but are not limited to the following:

Adaptation, Detection, Estimation, and Learning
- Distributed detection and estimation
- Distributed adaptation over networks
- Distributed learning over networks
- Distributed target tracking
- Bayesian learning; Bayesian signal processing
- Sequential learning over networks
- Decision making over networks
- Distributed dictionary learning
- Distributed game theoretic strategies
- Distributed information processing
- Graphical and kernel methods
- Consensus over network systems
- Optimization over network systems

Communications, Networking, and Sensing
- Distributed monitoring and sensing
- Signal processing for distributed communications and networking
- Signal processing for cooperative networking
- Signal processing for network security
- Optimal network signal processing and resource allocation

Modeling and Analysis
- Performance and bounds of methods
- Robustness and vulnerability
- Network modeling and identification

Modeling and Analysis (cont.)
- Simulations of networked information processing systems
- Social learning
- Bio-inspired network signal processing
- Epidemics and diffusion in populations

Imaging and Media Applications
- Image and video processing over networks
- Media cloud computing and communication
- Multimedia streaming and transport
- Social media computing and networking
- Signal processing for cyber-physical systems
- Wireless/mobile multimedia

Data Analysis
- Processing, analysis, and visualization of big data
- Signal and information processing for crowd computing
- Signal and information processing for the Internet of Things
- Emergence of behavior

Emerging topics and applications
- Emerging topics
- Applications in life sciences, ecology, energy, social networks, economic networks, finance, social sciences, smart grids, wireless health, robotics, transportation, and other areas of science and engineering

Editor-in-Chief: Petar M. Djurić, Stony Brook University (USA)

To submit a paper, go to: https://mc.manuscriptcentral.com/tsipn-ieee

IEEE COMMUNICATIONS SOCIETY

IEEE Signal Processing Society
http://www.signalprocessingsociety.org/

IEEE COMPUTER SOCIETY

$$R_{[\cdot][\cdot]} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \cdot \begin{pmatrix} 4 & 6 & 10 & 12 \\ 2 & 6 & 8 & 12 \\ 1 & 4 & 6 & 7 \\ 0 & 3 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 9 & -3 & 21 & -3 \\ 1 & 1 & 1 & 1 \\ 4 & -3 & 8 & 0 \\ 1 & 0 & 5 & -1 \end{pmatrix}.$$

■ $d_{[\cdot][\cdot]}$ is now formed by the elements of 2-D-DWT($S_{[\cdot][\cdot]}$), which are on the odd rows and odd columns

2) The rearranged matrix corresponds to $R_{[\cdot][\cdot]}$. The original signal is $S_{[\cdot][\cdot]} = A_{[\cdot][\cdot]}^{T} R_{[\cdot][\cdot]} A_{[\cdot][\cdot]}$.

NUMERICAL EXAMPLE
Let

$$S_{[\cdot][\cdot]} = \begin{pmatrix} 4 & 6 & 10 & 12 \\ 2 & 6 & 8 & 12 \\ 1 & 4 & 6 & 7 \\ 0 & 3 & 2 & 1 \end{pmatrix}, \quad h[\cdot] = \left\{ \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\}$$

and $g[\cdot] = \{1/\sqrt{2}, -1/\sqrt{2}\}$. To obtain 2-D-DWT ($S_{[\cdot][\cdot]}$), we need to calculate $R_{[\cdot][\cdot]} = A_{[\cdot][\cdot]} S_{[\cdot][\cdot]} A_{[\cdot][\cdot]}^{T}$, i.e., the boxed equation at the top of the page.

When rearranging the elements of this last matrix, we get:

$$\begin{pmatrix} 9 & 21 & -3 & -3 \\ 4 & 8 & -3 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 5 & 0 & -1 \end{pmatrix},$$

which corresponds to the 2-D-DWT ($S_{[\cdot][\cdot]}$). The inversion is directly based on the algorithm above.

CONCLUSIONS
This article offered and subsequently described tips on the 1-D and 2-D DWT calculations that were not yet been documented in literature in a practical way. The information presented herein can certainly help young students and experienced researchers in using this important tool for time-frequency signal analysis. I have developed a C/C++ source code that implements the DWT,

the IDWT, the DWPT, and the inverse DWPT (both 1-D and 2-D) containing examples of usage, and it is freely available. Please send your request to guido@ieee.org. Additional tips I wrote on DWT and DWTP can be found in [5].

AUTHOR
Rodrigo Capobianco Guido (guido@ieee.org) is a professor at São Paulo State University in São José do Rio Preto, Brazil.

REFERENCES
[1] P. Addison, J. Walker, and R. C. Guido, "Time-frequency analysis of biosignals: A wavelet transform overview," *IEEE Eng. Med. Biol. Mag.*, vol. 28, no. 5, pp. 14–29, 2009.

[2] S.-H. Chen, R. C. Guido, T.-K. Truong, and Y. Chang, "Improved voice activity detection algorithm using wavelet and support vector machine," *Comput. Speech Lang.*, vol. 24, no. 3, pp. 531–543, 2010.

[3] R. C. Guido, S. Barbon Jr., L. S. Vieira, F. L. Sanchez, C. D. Maciel, J. C. Pereira, P. R. Scalassara, and E. S. Fonseca, "Introduction to the discrete shapelet transform and a new paradigm: Joint time-frequency-shape analysis," in *Proc. IEEE Int. Symp. Circuits and Systems (IEEE ISCAS 2008)*, Seattle, WA, vol. 1, pp. 2893–2896.

[4] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge Press, 1997.

[5] R. C. Guido, "A note on a practical relationship between filters coefficients and the scaling and wavelet functions of the discrete wavelet transform," *Appl. Math. Lett.*, vol. 24, no. 7, pp. 1257–1259, 2011.

[SP]

[dates AHEAD]

Please send calendar submissions to:
Dates Ahead, c/o Jessica Barragué
IEEE Signal Processing Magazine
445 Hoes Lane
Piscataway, NJ 08855 USA
e-mail: j.barrague@ieee.org

2015

APRIL

14th IEEE International Conference on Information Processing in Sensor Networks (IPSN)
13–17 April, Seattle, Washington, United States.
General Chair: Suman Nath
URL: http://ipsn.acm.org/2015

First IEEE Conference on Network Softwarization (NetSoft)
13–17 April, London, United Kingdom.
General Cochairs: Prosper Chemouil
and George Pavlou
URL: http://sites.ieee.org/netsoft/

12th IEEE International Symposium on Biomedical Imaging (ISBI)
16–19 April, Brooklyn, New York, United States.
General Chairs: Elsa Angelini and
Jelena Kovačević
URL: http://biomedicalimaging.org/2015/

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
19–24 April, Brisbane, Australia.
General Cochairs: Vaughan Clarkson
and Jonathan Manton
URL: http://icassp2015.org/

MAY

31st Picture Coding Symposium (PCS)
31 May–3 June, Cairns, Australia.
General Chairs: David Taubman
and Mark Pickering
URL: http://www.pcs2015.org

JUNE

IEEE Signal Processing Society Summer School on Biomedical Image Processing and Analysis (SSBIPA)
13–19 June, Dubrovnik, Croatia.
General Chair: Sven Lončarić
URL: https://sites.google.com/site/ssbipa2015/

Digital Object Identifier 10.1109/MSP.2014.2387614
Date of publication: 6 April 2015

Third IEEE International Workshop on Compressed Sensing Theory and Its Applications to Radar, Sonar, and Remote Sensing (CoSeRa)
22–24 June, Pisa, Italy.
General Chairs: Fulvio Gini and Joachim Ender
URL: http://www.cosera2015.iet.unipi.it/

IEEE Signal Processing Society Summer School on Foundations and Advances in Stochastic Filtering (FASF)
22–26 June, Barcelona, Spain.
Organizers: Pau Closas and Joaquín Míguez
URL: http://fasf2015.cttc.cat/

16th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)
28 June–1 July, Stockholm, Sweden.
General Chairs: Joakim Jaldén and
Björn Ottersten
URL: http://www.spawc2015.org/

IEEE International Conference on Multimedia and Expo (ICME)
29 June–3 July, Turin, Italy.
General Chairs: Enrico Magli,
Stefano Tubaro, and Anthony Vetro
URL: http://www.icme2015.ieee-icme.org/index.php

JULY

Third IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)
12–15 July, Chengdu, China.
General Chairs: Yingbo Hua and Dezhong Yao
URL: http://www.chinasip2015.org/

AUGUST

IEEE Signal Processing and SP Education Workshop (SPW)
9–12 August, Salt Lake City, Utah, United States.
General Chair: Todd Moon
URL: http://spw2015.coe.utah.edu/

12th IEEE International Conference on Advanced Video- and Signal-Based Surveillance (AVSS)
25–28 August, Karlsruhe, Germany.
General Chairs: Jürgen Beyerer
and Rainer Stiefelhagen
URL: http://avss2015.org

2015 23rd European Signal Processing Conference (EUSIPCO)
31 August–4 September, Nice, France.
General Chairs: Jean-Luc Dugelay and
Dirk Slock
URL: http://www.eusipco2015.org

SEPTEMBER

IEEE Signal Processing Society Italy Chapter Summer School on Signal Processing (S3P)
7–11 September, Brescia, Italy.

Sensor Signal Processing for Defence (SSPD)
9–10 September, Edinburgh, Scotland, United Kingdom.
General Chairs: Mike Davies, Jonathon Chambers, and Paul Thomas
URL: http://www.sspdconference.org

IEEE International Conference on Image Processing (ICIP)
28 September–1 October, Quebec City, Quebec, Canada.
URL: http://www.icip2015.org/

OCTOBER

IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB)
4–7 October, Montreal, Canada.
URL: http://www.icuwb2015.org/index.html

IEEE Workshop on Signal Processing Systems (SiPS)
14–16 October, Hangzhou, China.
General Chairs: Chaitali Chakrabarti
and Nam Ling
URL: http://www.sips2015.org/

IEEE International Workshop on Multimedia Signal Processing (MMSP)
19–21 October, Xiamen, China.
General Chairs: Xiao-Ping Zhang,
Oscar C. Au, and Jonathan Li
URL: http://www.mmsp2015.org/

NOVEMBER

Seventh IEEE International Workshop on Information Forensics and Security (WIFS)
16–19 November, Rome, Italy.
General Chair: Patrizio Campis
URL: http://www.wifs2015.org/

DECEMBER

IEEE 6th International Workshop on Computational Advances in Multisensor Adaptive Processing (CAMSAP)
13–16 December, Cancun, Mexico.

IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)
13–17 December, Scottsdale, Arizona, United States.
URL: http://www.asru2015.org/

IEEE Global Conference on Signal and Information Processing (GlobalSIP)
14–16 December, Orlando, Florida, United States.
General Chairs: José M.F. Moura
and Dapeng Oliver Wu

SP
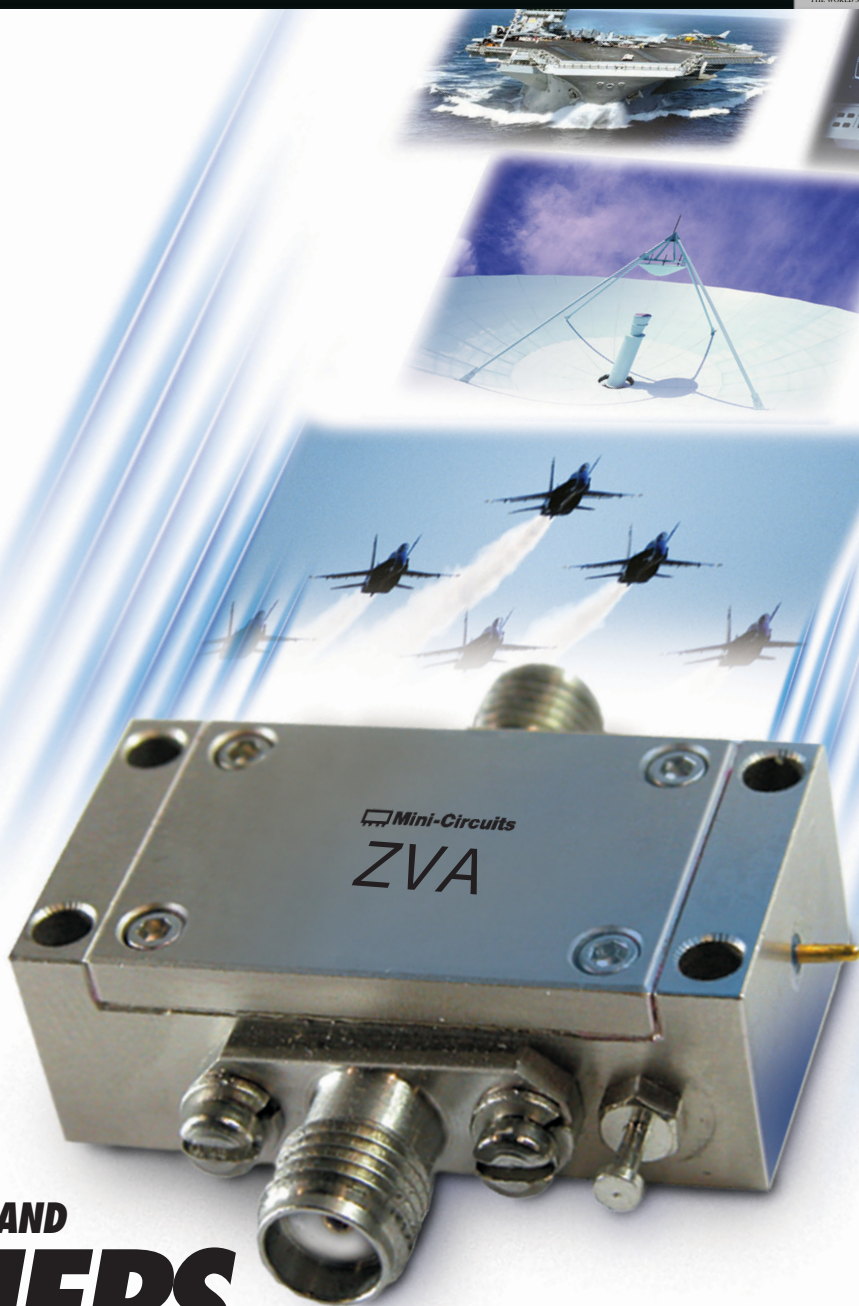
SUPER ULTRA WIDEBAND
AMPLIFIERS

up to +27 dBm output... 0.1 to 21 GHz

$^{\$}$845 *from* *ea.*

Ultra wide coverage and super flat gain make our ZVA family ideal for ECM, instrumentation, and test systems. With output power up to 0.5 Watts, they're simply some of the most usable amplifiers you'll find, for a wide range of applications and architectures!

All of our ZVA models are unconditionally stable, ruggedly constructed, and able to withstand open or short circuits at full output. For more details, from data sheets to environmental ratings, pricing, and real-time availability, just go to minicircuits.com!
All models IN STOCK!

◔ *RoHS compliant*

Electrical Specifications (-55 to +85°C base plate temperature)

Model	Frequency (GHz)	Gain (dB)	P1dB (dBm)	IP3 (dBm)	NF (dB)	Price $ * (Qty. 1-9)
NEW						
ZVA-183WX+	0.1-18	28±2	27	35	3.0	1345.00
ZVA-183X+	0.7-18	26±1	24	33	3.0	845.00
ZVA-213X+	0.8-21	26±2	24	33	3.0	945.00

* Heat sink must be provided to limit base plate temperature.To order with heat sink, remove "X" from model number and add $50 to price.

Wideband Performance
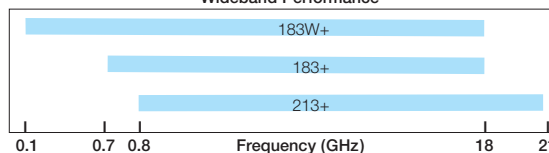
183W+	
183+	
213+	

0.1   0.7 0.8   **Frequency (GHz)**   18   21

Mini-Circuits®

MathWorks®
Accelerating the pace of engineering and science

Parlez-vous MATLAB?

Over one million people around the world speak MATLAB. Engineers and scientists in every field from aerospace and semiconductors to biotech, financial services, and earth and ocean sciences use it to express their ideas.
Do you speak MATLAB?
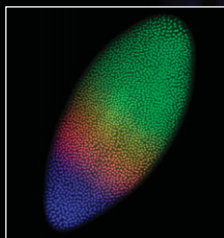
Quantitative high-throughput gene expression imaging using data from FlyEx Database.

This example available at mathworks.com/ltc

Image from FlyEx Database. Used by permission: http://flyex.ams.sunysb.edu/flyex and http://urchin.spbcas.ru/flyex ©2015 The MathWorks, Inc.

MATLAB®

The language of technical computing.

IEEE SIGNAL PROCESSING SOCIETY

CONTENT GAZETTE

[ISSN 2167-5023]

MAY 2015

IEEE Signal Processing Society
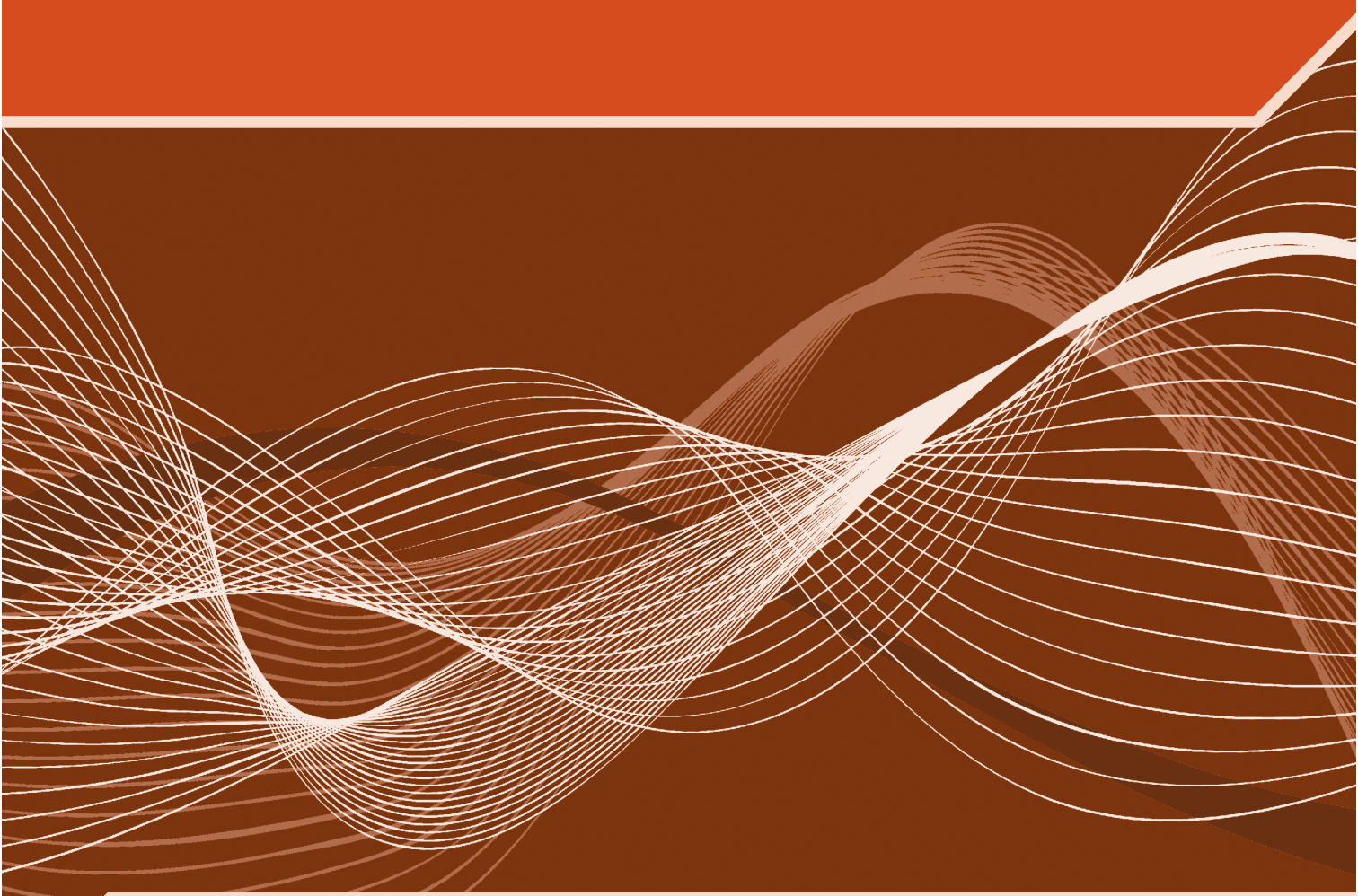
IEEE

ASRU 2015

IEEE Automatic Speech Recognition and Understanding Workshop

December 13-17, 2015 Scottsdale, Arizona, USA

http://asru2015.org

Call for Papers

The fourteenth biannual IEEE workshop on Automatic Speech Recognition and Understanding (ASRU) will be held on December 13-17, 2015 in Scottsdale, Arizona - USA. The ASRU workshop meets every two years and has a tradition of bringing together researchers from academia and industry in an intimate and collegial setting to discuss problems of common interest in **automatic speech recognition, understanding, and related fields of research**.

Topics and focus

Authors are encouraged to submit contributions in all areas of spoken language processing, with emphasis placed on the following topics:

- Automatic speech recognition
- Spoken language understanding
- Speech-to-text systems
- Spoken dialog systems
- Multilingual language processing
- Robustness in automatic speech recognition
- Spoken document retrieval
- Speech-to-speech translation
- Text-to-speech systems
- Spontaneous speech processing
- Speech summarization
- New applications of automatic speech recognition

Format

The workshop features one keynote and one or two invited talks a day. Regular papers are presented as posters. See http://asru2015.org for formatting guidelines. ASRU 2015 will also include challenge tasks, panel discussions and demo sessions.

Paper Submission

Prospective authors are invited to submit full-length, 4-6 page papers, including figures, plus 1-2 additional pages for references only. All papers will be handled and reviewed electronically.

Schedule

Paper due date: **Friday July 10, 2015**
Paper Notification: Friday Sept 11, 2015
Registration opens: Friday Sept 11, 2015
Demo/toolkit deadline: Friday Sept 25, 2015
Paper Camera ready version due: Friday Oct 2, 2015
Demo/toolkit notification date: Friday Oct 9, 2015
Author and early registration end: Friday Oct 23, 2015
Demo/toolkit camera ready version due: Monday Oct 26, 2015
Workshop: Dec 13-17, 2015

IEEE TRANSACTIONS ON
SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society ®

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed

MEDLINE
U.S. National Library of Medicine

REGULAR PAPERS

◈IEEE

IEEE TRANSACTIONS ON

SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society ®

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed

MEDLINE
U.S. National Library of Medicine

REGULAR PAPERS

◆IEEE

IEEE TRANSACTIONS ON

SIGNAL AND INFORMATION PROCESSING OVER NETWORKS

Now accepting paper submissions

The new *IEEE Transactions on Signal and Information Processing over Networks* publishes high-quality papers that extend the classical notions of processing of signals defined over vector spaces (e.g. time and space) to processing of signals and information (data) defined over networks, potentially dynamically varying. In signal processing over networks, the topology of the network may define structural relationships in the data, or may constrain processing of the data. Topics of interest include, but are not limited to the following:

Adaptation, Detection, Estimation, and Learning
- Distributed detection and estimation
- Distributed adaptation over networks
- Distributed learning over networks
- Distributed target tracking
- Bayesian learning; Bayesian signal processing
- Sequential learning over networks
- Decision making over networks
- Distributed dictionary learning
- Distributed game theoretic strategies
- Distributed information processing
- Graphical and kernel methods
- Consensus over network systems
- Optimization over network systems

Communications, Networking, and Sensing
- Distributed monitoring and sensing
- Signal processing for distributed communications and networking
- Signal processing for cooperative networking
- Signal processing for network security
- Optimal network signal processing and resource allocation

Modeling and Analysis
- Performance and bounds of methods
- Robustness and vulnerability
- Network modeling and identification

Modeling and Analysis (cont.)
- Simulations of networked information processing systems
- Social learning
- Bio-inspired network signal processing
- Epidemics and diffusion in populations

Imaging and Media Applications
- Image and video processing over networks
- Media cloud computing and communication
- Multimedia streaming and transport
- Social media computing and networking
- Signal processing for cyber-physical systems
- Wireless/mobile multimedia

Data Analysis
- Processing, analysis, and visualization of big data
- Signal and information processing for crowd computing
- Signal and information processing for the Internet of Things
- Emergence of behavior

Emerging topics and applications
- Emerging topics
- Applications in life sciences, ecology, energy, social networks, economic networks, finance, social sciences, smart grids, wireless health, robotics, transportation, and other areas of science and engineering

Editor-in-Chief: Petar M. Djurić, Stony Brook University (USA)
To submit a paper, go to: https://mc.manuscriptcentral.com/tsipn-ieee

IEEE COMMUNICATIONS SOCIETY

IEEE Signal Processing Society

IEEE COMPUTER SOCIETY

Announcing Signal Processing's newest publication:
Transactions on Signal and Information Processing over Networks (T-SIPN)

>>We are accepting submissions: please submit a manuscript here<<

There has been an explosion of research in network systems of various types, including physical, engineered, biological and social systems. Its aim is to find answers to fundamental questions about the systems and with them be able to understand, predict, and control them better. To that end, a core area of work is signal and information processing over networks.

Network systems represent a growing research field encompassing numerous disciplines in science and engineering. Their complexity is reflected in the diversity and the interconnectivity of their elements, which have the capacity to adapt and learn from experience. Applications of network systems are wide and include communications (wireless sensor networks, peer-to-peer networks, pervasive mobile networks, the Internet of Things), the electric power grid, biology, the Internet, the stock market, ecology, and in animal and human societies.

The Transactions on Signal and Information Processing over Networks (T-SIPN) publishes timely peer-reviewed technical articles on advances in the theory, methods, and algorithms for signal and information processing, inference, and learning in network systems. The following core topics define the scope of the Transaction:

Adaptation, Detection, Estimation, and Learning (ADEL)

- o Distributed detection and estimation   (ADEL-DDE)
- o Distributed adaptation over networks (ADEL-DAN)
- o Distributed learning over networks (ADEL-DLN)
- o Distributed target tracking (ADEL-DTT)
- o Bayesian learning; Bayesian signal processing (ADEL-BLSP)
- o Sequential learning over networks (ADEL-SLN)
- o Decision making over networks (ADEL-DMN)
- o Distributed dictionary learning (ADEL-DDL)
- o Distributed game theoretic strategies (ADEL-DGTS)
- o Distributed information processing (ADEL-DIP)
- o Graphical and kernel methods (ADEL-GKM)
- o Consensus over network systems (ADEL-CNS)
- o Optimization over network systems (ADEL-ONS)

Communications, Networking, and Sensing (CNS)

- o Distributed monitoring and sensing (CNS-DMS)
- o Signal processing for distributed communications and networking (CNS-SPDCN)
- o Signal processing for cooperative networking (CNS-SPCN)
- o Signal processing for network security (CNS-SPNS)
- o Optimal network signal processing and resource allocation (CNS-NSPRA)

IEEE/ACM TRANSACTIONS ON
AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society ®

(acm) Association for Computing Machinery

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed

MEDLINE
U.S. National Library of Medicine

SPECIAL SECTION ON CONTINUOUS SPACE AND RELATED METHODS IN NATURAL LANGUAGE PROCESSING

◆IEEE

IEEE/ACM TRANSACTIONS ON
AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society ®

acm Association for Computing Machinery

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed

MEDLINE
U.S. National Library of Medicine

REGULAR PAPERS

◆IEEE

The Ninth IEEE Sensor Array and Multichannel Signal Processing Workshop

IEEE SAM 2016
Rio de Janeiro
Brazil

10th-13th July 2016, Rio de Janeiro, Brazil

IEEE

IEEE Signal Processing Society

Call for Papers

General Chairs

Rodrigo C. de Lamare,
PUC-Rio, Brazil and University of York, United Kingdom

Martin Haardt,
TU Ilmenau, Germany

Technical Chairs

Aleksandar Dogandzic,
Iowa State University, USA

Vítor Nascimento,
University of São Paulo, Brazil

Special Sessions Chair

Cédric Richard,
University of Nice, France

Publicity Chair

Maria Sabrina Greco,
Universityof Pisa, Italy

Important Dates

Special Session Proposals
29thJanuary , 2016

Submission of Papers
26thFebruary, 2016

Notification of Acceptance
29thApril , 2016

Final Manuscript Submission
16th May, 2016

Advance Registration
16th May, 2016

Technical Program
The SAM Workshop is an important IEEE Signal Processing Society event dedicated to sensor array and multichannel signal processing. The organizing committee invites the international community to contribute with state-of-the-art developments in the field. SAM 2016 will feature plenary talks by leading researchers in the field as well as poster and oral sessions with presentations by the participants.

Welcome to Rio de Janeiro! – The workshop will be held at the Pontifical Catholic University of Rio de Janeiro, located in Gávea, in a superb area surrounded by beaches, mountains and the Tijuca National Forest, the world's largest urban forest. Rio de Janeiro is a world renowned city for its culture, beautiful landscapes, numerous tourist attractions and international cuisine. The workshop will take place during the first half of July about a month before the 2016 Summer Olympic Games when Rio will offer plenty of cultural activities and festivities, which will make SAM 2016 a memorable experience.

Research Areas
Authors are invited to submit contributions in the following areas:
- Adaptive beamforming
- Array processing for biomedical applications
- Array processing for communications
- Blind source separation and channel identification
- Computational and optimization techniques
- Compressive sensing and sparsity-based signal processing
- Detection and estimation
- Direction-of-arrival estimation
- Distributed and adaptive signal processing
- Intelligentsystems and knowledge-based signal processing
- Microphone and loudspeaker array applications
- MIMO radar
- Multi-antenna systems: multiuser MMO, massive MIMO and space-time coding
- Multi-channel imaging and hyperspectral processing
- Multi-sensor processing for smart grid and energy
- Non-Gaussian, nonlinear, and non-stationary models
- Performance evaluations with experimental data
- Radar and sonar array processing
- Sensor networks
- Source Localization, Classification and Tracking
- Synthetic aperture techniques
- Space-time adaptive processing
- Statistical modelling for sensor arrays
- Waveform diverse sensors and systems

Submission of papers – Full-length four-page papers will be accepted only electronically.

Special session proposals – They should be submitted by e-mail to the Technical Program Chairs and the Special Sessions Chair and include a topical title, rationale, session outline, contact information, and list of invited speakers.

IEEE TRANSACTIONS ON

IMAGE PROCESSING

PAPERS

IEEE TRANSACTIONS ON
IMAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed   MEDLINE
U.S. National Library of Medicine

IEEE

IEEE TRANSACTIONS ON
COMPUTATIONAL IMAGING

NEW!

The new IEEE Transactions on Computational Imaging seeks original manuscripts for publication. This new journal will publish research results where computation plays an integral role in the image formation process. All areas of computational imaging are appropriate, ranging from the principles and theory of computational imaging, to modeling paradigms for computational imaging, to image formation methods, to the latest innovative computational imaging system designs. Topics of interest include, but are not limited to the following:

Imaging Models and Representation

- Statistical-model based methods
- System and image prior models
- Noise models
- Graphical and tree-based models
- Perceptual models

Computational Sensing

- Coded source methods
- Structured light
- Coded aperture methods
- Compressed sensing
- Light-field sensing
- Plenoptic imaging
- Hardware and software systems

Computational Image Creation

- Sparsity-based methods
- Statistically-based inversion methods, Bayesian regularization
- Super-resolution, multi-image fusion
- Learning-based methods, Dictionary-based methods
- Optimization-based methods; proximal iterative methods, ADMM

Computational Photography

- Non-classical image capture, Generalized illumination
- Time-of-flight imaging
- High dynamic range imaging
- Focal stacks

Computational Consumer Imaging

- Cell phone imaging
- Camera-array systems
- Depth cameras

Computational Acoustic Imaging

- Multi-static ultrasound imaging
- Photo-acoustic imaging
- Acoustic tomography

Computational Microscopic Imaging

- Holographic microscopy
- Quantitative phase imaging
- Multi-illumination microscopy
- Lensless microscopy

Tomographic Imaging

- X-ray CT
- PET
- SPECT

Magnetic Resonance Imaging

- Diffusion tensor imaging
- Fast acquisition

Radar Imaging

- Synthetic aperture imaging
- Inverse synthetic imaging
- Terahertz imaging

Geophysical Imaging

- Multi-spectral imaging
- Ground penetrating radar
- Seismic tomography

Multi-spectral Imaging

- Multi-spectral imaging
- Hyper-spectral imaging
- Spectroscopic imaging

Editor-in-Chief: W. Clem Karl, Boston University.
To submit a paper go to: https://mc.manuscriptcentral.com/tci-ieee

IEEE
Signal Processing Society

IEEE International Conference on Image Processing

Phoenix Convention Center

September 25 - 28, 2016   Phoenix, Arizona

General Chair
Lina Karam
Arizona State University

General Co-Chair
Aggelos Katsaggelos
Northwestern University

Technical Program Chairs
Fernando Pereira
Instituto Superior Técnico
Gaurav Sharma
University of Rochester

Innovation Program Chairs
Haohong Wang
TCL Research America
Jeff Bier
BDTI & Embedded Vision Alliance

Finance Chair
Sohail Dianat
Rochester Institute of Technology

Plenary Chairs
Michael Marcellin
University of Arizona
Sethuraman Panchanathan
Arizona State University

Special Sessions Chairs
Dinei Florencio
Microsoft Research
Chaker Larabi
Poitiers University
Zhou Wang
University of Waterloo

Tutorials Chairs
Ghassan AlRegib
Georgia Tech
Rony Ferzli
Intel

Publicity Chair
Michel Sarkis
Qualcomm Technologies Inc.

Awards Chairs
Vivek Goyal
Boston University
Ivana Tosic
Ricoh Innovations

Exhibits Chair
David Frakes
Arizona State University &
Google

Publication Chairs
Patrick Le Callet
Nantes University
Baoxin Li
Arizona State University

Local Arrangement Chairs
Jorge Caviedes
Intel
Pavan Turaga
Arizona State University

Registration Chair
Ricardo De Queiroz
Universidade de Brasilia

Conference Management
Conference Management Services

The 23rd IEEE International Conference on Image Processing (ICIP) will be held in the Phoenix Convention Centre, Phoenix, Arizona, USA, on September 25 - 28, 2016. ICIP is the world's largest and most comprehensive technical conference focused on image and video processing and computer vision. In addition to the Technical Program, ICIP 2016 will feature **an Innovation Program focused on innovative vision technologies and fostering innovation, entrepreneurship, and networking**. The conference will feature world-class speakers, tutorials, exhibits, and a vision technology showcase.

Topics in the ICIP 2016 Technical Program include but are not limited to the following:

Filtering, Transforms, Multi-Resolution Processing
Restoration, Enhancement, Super-Resolution
Computer Vision Algorithms and Technologies
Compression, Transmission, Storage, Retrieval
Computational Imaging
Color and Multispectral Processing
Multi-View and Stereoscopic Processing
Multi-Temporal and Spatio-Temporal Processing
Video Processing and Analytics
Authentication and Biometrics

Biological and Perceptual-based Processing
Visual Quality Assessment
Scanning, Display, and Printing
Document and Synthetic Visual Processing
Applications to various fields (e.g., biomedical, Advanced Driving Assist Systems, assistive living, security, learning, health and environmental monitoring, manufacturing, consumer electronics)

The ICIP 2016 innovation program will feature a **vision technology showcase** of state-of-the-art vision technologies, innovation challenges, talks by innovation leaders and entrepreneurs, tutorials, and networking.

Paper Submission: Prospective authors are invited to submit full-length papers at the conference website, with up to four pages for technical content including figures and references, and with one additional optional 5th page for references only. Submission instructions, templates for the required paper format, and information on "no show" policy are available at **www.icip2016.com**.

Tutorials and Special Sessions Proposals: Tutorials will be held on September 25, 2016. Tutorial proposals should be submitted to *tutorials@icip2016.com* and must include title, outline, contact information, biography and selected publications for the presenter(s), and a description of the tutorial and material to be distributed to participants. Special Sessions proposals should be submitted to *specialsessions@icip2016.com* and must include a topical title, rationale, session outline, contact information, and a list of invited papers. For detailed submission guidelines, please refer the ICIP 2016 website at **www.icip2016.com**.

> **Important Deadlines:**
> Special Session and Tutorial Proposals: November 16, 2015
> Notification of Special Session and Tutorial Acceptance: December 18, 2015
>
> Paper Submissions: January 25, 2016
> Notification of Paper Acceptance: April 30, 2016
>
> Visual Technology Innovator Award Nomination: March 30, 2016
>
> Revised Paper Upload Deadline: May 30, 2016
> Authors' Registration Deadline: May 30, 2016

http://www.facebook.com/icip2016
https://twitter.com/icip2016/
https://www.linkedin.com/groups/ICIP-2016-6940658

IEEE TRANSACTIONS ON

INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

PAPERS

IEEE GlobalSIP'15–Call for Papers

2015 IEEE Global Conference on Signal and Information Processing – Orlando Florida
General Chairs: Jose Moura and Dapeng Oliver Wu
Technical Program Chairs: Mihaela van der Schaar, Xiaodong Wang, and Hsiao-Chun Wu

The IEEE Global Conference on Signal and Information Processing (GlobalSIP) is a recently launched flagship conference of the *IEEE Signal Processing Society*. GlobalSIP' 15 will be held in Orlando, Florida, USA, December 14-16, 2015. The conference will focus broadly on signal and information processing with an emphasis on up-and-coming signal processing themes. The conference will feature world-class speakers, tutorials, exhibits, and technical sessions consisting of poster or oral presentations. GlobalSIP' 15 technical program will be comprised of a main program and several co-located symposia on special topics. Technical paper submissions are solicited in the interest topics which may include, but are not limited to:

- Signal processing in communications and networks, including green communication and Signal processing in optical communication
- Image and video processing
- Selective topics in speech and language processing
- Signal processing in security applications
- Signal processing in finance
- Signal processing in energy and power systems
- Signal processing in genomics and bioengineering (physiological, pharmacological and behavioral)
- Signal processing for social media networks
- Neural signal processing
- Seismic signal processing
- Selective topics in statistical signal processing
- Graph-theoretic signal processing
- Machine learning
- Compressed sensing, sparsity analysis, and applications
- Big data processing, heterogeneous information processing and informatics
- Human machine interfaces
- Multimedia transmission, indexing and retrieval, and playback challenges
- Hardware and real-time implementations
- Other novel and significant Applications of selected areas of signal processing

Submission of Papers: Prospective authors are invited to submit full-length papers, with up to four pages for technical content including figures and possible references, and with one additional optional 5th page containing only references. Manuscripts should be original (not submitted/published anywhere else) and written in accordance with the standard IEEE double-column paper template. All paper submissions should be carried out through EDAS system (http://edas.info). A selection of best papers and best student papers will be made by the GlobalSIP 2015 best paper award committee upon recommendations from Technical Committees.

Notice: The IEEE Signal Processing Society enforces a "no-show" policy. Any accepted paper included in the final program is expected to have at least one author or qualified proxy attend and present the paper at the conference. Authors of the accepted papers included in the final program who do not attend the conference will be subscribed to a "No-Show List", compiled by the Society. The "no-show" papers will not be published by IEEE on IEEEXplore or other public access forums, but these papers will be distributed as part of the on-site electronic proceedings and the copyright of these papers will belong to the IEEE.

Timeline for paper submission:

May 15, 2015:	Paper submission deadline
June 30, 2015:	Review results announced
September 5, 2015:	Camera-ready papers due

WIFS 2015

7th IEEE International Workshop on Information Forensics and Security

Rome, Italy, November 16-19, 2015

General Chair
Patrizio Campisi,
Roma Tre University, Italy

General co-Chair
Nasir Memon
New York University, USA

Technical Program Chairs
Fernando Pérez-González
University of Vigo, Spain
Slava Voloshynovskiy
University of Geneva, Switzerland

Tutorials Chairs
Tanya T. Ignatenko
TUE, The Netherlands
Zekeriya Erkin
TUDelft, The Netherlands

Demo Session Chair
Samson Cheung
University of Kentucky, USA

Publications Chair
Emanuele Maiorana
Roma Tre University, Italy

Finance Chair
Stefano Tubaro
Polytechnic Univ. of Milan, Italy

Publicity Chairs
Marco Carli
Roma Tre University, Italy
Anderson Rocha
University of Campinas, Brasil

Industry Liaison
Dinei Florencio
Microsoft Research, USA
Hervé Chabanne
Morpho, France
Tomas Filler
Digimarc, USA

European Liaison
Jana Dittman
Univ. of Magdeburg, Germany

American Liaison
Arun Ross
Michigan State University, USA

Asian Liaison
Alex C. Kot
NTU, Singapore

Local Arrangements Chair
Federica Battisti
Roma Tre University, Italy

The **IEEE International Workshop on Information Forensics and Security (WIFS)** is the annual flagship workshop organized by the IEEE Information Forensics and Security (IFS) Technical Committee. Its major goal is to bring together researchers in the field to foster ideas exchange and to allow cross-fertilization among researchers working in the different areas of information security. The 7th edition of WIFS will be held in Rome, Italy, hosted by **Roma Tre University**, from November 16, to November 19, 2015. The conference will feature three keynote lectures, up to four tutorials, lectures and poster sessions, and also demo and ongoing works sessions. Topics of interest include, but are not limited to:

- Amonymization and Data Privacy
- Cryptography for multimedia
- Biometrics
- Surveillance
- Forensic Analysis
- Hardware Security
- Information Theoretic Security
- Multimedia Hashing
- Network Security
- Adversarial Signal Processing
- Communication and Physical-Layer Security
- Steganography and covert communications
- Usability and Human Factors
- Watermarking and Data Hiding

Submission of papers: Prospective authors are invited to submit full-length, six-page papers, including figures and references. The WIFS Technical Program Committee will select the best submitted papers to be presented at WIFS 2015. Accepted papers will be scheduled in lecture tracks or in poster sessions. Authors are required to present their papers at the conference.

Tutorial proposals: Up to four tutorials will be scheduled for the first day of the conference, Monday November 16, 2015. Prospective tutorial contributors are encouraged to submit a tutorial proposal with the tutorial title, the presenters' name, affiliation, and brief CV, along with the detailed structure of the tutorial, to the Tutorials Chairs at tutorials@wifs2015.org.

Demo and ongoing works proposals: The Demo session is open to both academic researchers and industrial exhibitors. An ongoing works session will be organized to give researchers the opportunity to present their latest activities to the general audience. Formal proposals have to be submitted to the Demo Session Chair at demo@wifs2015.org.

Submission of SPL and TIFS papers: Authors of IEEE Signal Processing Letters (SPL) and IEEE Transactions on Information Forensics and Security (TIFS) papers will be given the opportunity to present their work at WIFS 2015, subject to space availability and approval by the WIFS Technical Program Chairs. Proposals have to be submitted to the Technical Program Chairs at tpc@wifs2015.org.

Important dates

For further details, please visit the WIFS 2015 conference website: **http://www.wifs2015.org**

IEEE TRANSACTIONS ON

MULTIMEDIA

A PUBLICATION OF
THE IEEE CIRCUITS AND SYSTEMS SOCIETY
THE IEEE SIGNAL PROCESSING SOCIETY
THE IEEE COMMUNICATIONS SOCIETY
THE IEEE COMPUTER SOCIETY

http://www.signalprocessingsociety.org/tmm/

IEEE JOURNAL OF
SELECTED TOPICS IN SIGNAL PROCESSING

MLSP2015

IEEE International Workshop on
MACHINE LEARNING FOR SIGNAL PROCESSING

SEPTEMBER 17-20, 2015   BOSTON, MASSACHUSSETTS, USA

MLSP2015.CONWIZ.DK

Organizing committee

General Chair	**Deniz Erdoğmuş** *Northeastern University*
Program Chairs	**Murat Akçakaya** *University of Pittsburgh*
	Serdar Kozat *Bilkent University*
Data Competition Chair	**Vince Calhoun** *University of New Mexico*
Special Session Chair	**Catherine Huang** *Intel Labs*
Publicity Chair	**Kostas Diamantaras** *TEI of Thessaloniki*
Publication Chair	**Jan Larsen** *Technical University of Denmark*

Important Dates and Deadlines

Special session proposals	April 12, 2015
Paper submissions	May 17, 2015
Decision notifications	June 21, 2015
Camera-ready papers due	June 28, 2015
Advance registration	August 2, 2015

Supported by

IEEE Signal Processing Society

Northeastern University

CALL FOR PAPERS

The 25th MLSP workshop in the series of workshops organized by the IEEE Signal Processing Society MLSP Technical Committee will present the most recent and exciting advances in machine learning for signal processing through keynote talks, tutorials, as well as special and regular single-track sessions. Prospective authors are invited to submit papers on relevant algorithms and applications including, but not limited to:

- Learning theory and techniques
- Graphical models and kernel methods
- Data-driven adaptive systems and models
- Pattern recognition and classification
- Distributed, Bayesian, subspace/manifold/sparsity-aware learning
- Multiset data analysis and multimodal data fusion
- Perceptual signal processing in audio, image and video
- Cognitive information processing
- Multichannel adaptive and nonlinear signal processing
- Applications, including: speech & audio, image & video, music, biomedical signals & images, communications, bioinformatics, biometrics, computational intelligence, genomic signals & sequences, social networks, games, smart grid, security & privacy

Data Analysis and Signal Processing Competition is organized in conjunction with the workshop. The goal of the competition is to advance the current state-of-the-art in theoretical and practical aspects of signal processing domains.

MLSP 2015 seeks proposals for **Special Sessions** that will address research in emerging or interdisciplinary areas of particular interest, not covered already by traditional MLSP sessions. Please submit proposals to the Special Session Chair.

The **MLSP Best Student Paper Award** will be granted to the best paper for which a student is the principal author and presenter.

Prospective authors are invited to submit a double column paper of up to six pages using the electronic submission procedure at http://mlsp2015.conwiz.dk. Accepted papers will be published on a password-protected website that will be available during the workshop. The presented papers will be published in and indexed by IEEE Xplore.

Please refer to the workshop website for more details.

Signal and Information Processing: The Heartbeat of a Smart Society

Shanghai, China

ICASSP·2016

41st IEEE International Conference on Acoustics, Speech and Signal Processing, 20-25 March 2016
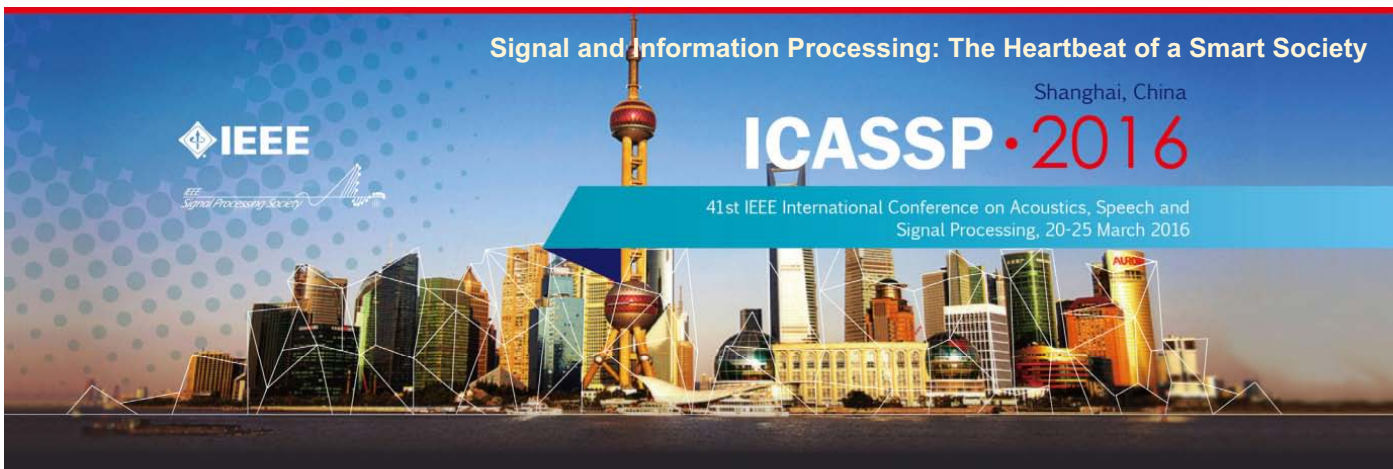
General Chairs
Zhi Ding, Univ. of California, Davis, USA
Zhi-Quan Luo, Univ. of Minnesota, USA
Wenjun Zhang, Shanghai Jiao Tong Univ., China

Technical Program Chairs
P. C. Ching, Chinese Univ. of Hong Kong, Hong Kong
Dominic K.C. Ho, Univ. of Missouri, USA

Finance Chairs
Shuguang Cui, Texas A&M Univ., USA
Rong Xie, Shanghai Jiao Tong Univ., China

Plenaries Chairs
Zhi-Pei Liang, UIUC, USA
Björn Ottersten, Univ. of Luxembourg, Luxembourg

Special Sessions Chairs
Tim Davidson, McMaster Univ., Canada
Jianguo Huang, Northwestern Polytech. Univ., China

Tutorials Chairs
Jian Li, Univ. of Florida, USA
Jose Principe, Univ. of Florida, USA

Student Session Chair
Wei Zhang, Univ. of New South Wales, Australia

Registration Chairs
Tongtong Li, Michigan State Univ., USA
Xiaojun Yuan, ShanghaiTech Univ., China

Publicity Chairs
Xiaokang Yang, Shanghai Jiao Tong Univ., China
Mounir Ghogho, Leeds Univ., UK
Ignacio Santamaria, Univ. of Cantabria, Spain

Publication Chairs
Min Dong, Univ. of Ontario Inst. of Tech., Canada
Thomas Fang Zheng, Tsinghua Univ., China

Industrial & Exhibit Chairs
Li Deng, Microsoft, USA
Jinyu Li, Microsoft, USA
Cathy Wicks, Texas Instruments, USA

Local Arrangement Chairs
Ning Liu, Shanghai Jiao Tong Univ., China
Meixia Tao, Shanghai Jiao Tong Univ., China

Webmaster
Yi Xu, Shanghai Jiao Tong Univ., China

Workshop Chairs
Jianguo Huang, Northwestern Polytech. Univ., China
Jiwu Huang, Sun Yat-sen Univ., China

ICASSP2016: Signal and information processing is the driving heartbeat in the development of technologies that enrich our lives and advance our society. The 41st International Conference on Acoustics, Speech, and Signal Processing (ICASSP) will be held in the Shanghai International Convention Center, Shanghai, China between March 20 and 25, 2016. The conference provides, both for researchers and developers, an engaging forum to exchange ideas and propel new developments in this field. The 2016 conference will showcase world-class presentations by internationally renowned speakers and will facilitate a fantastic opportunity to network with like-minded professionals from around the world. Topics include but are not limited to:

Audio and acoustic signal processing	Multimedia signal processing
Bio-imaging and biomedical signal processing	Sensor array & multichannel signal processing
Signal processing education	Design & implementation of signal processing systems
Speech processing	Signal processing for communications & networking
Industry technology tracks	Image, video & multidimensional signal processing
Information forensics and security	Signal processing theory & methods
Machine learning for signal processing	Spoken language processing
Signal processing for Big Data	Signal processing for the Internet of Things

Shanghai: Shanghai is the most populous city in China and one of the most populous cities in the world. A global city, Shanghai exerts influence over global commerce, finance, culture, art, fashion, research and entertainment. The city is located in the middle portion of the Chinese coast, and sits at the mouth of the Yangtze River. The city is a tourist destination renowned for its historical landmarks, such as the Bund and City God Temple, and its modern and ever-expanding Pudong skyline including the Oriental Pearl Tower. Today, Shanghai is the largest center of commerce and finance in mainland China, and has been described as the "showpiece" of the world's fastest-growing major economy.

Submission of Papers: Prospective authors are invited to submit full-length papers, with up to four pages for technical content including figures and possible references, and with one additional optional 5th page containing only references. A selection of best student papers will be made by the ICASSP 2016 committee upon recommendations from the Technical Committees.

Tutorial and Special Session Proposals: Tutorials will be held on March 20 and 21, 2016. Tutorial proposals must include title, outline, contact information, biography and selected publications for the presenter(s), and a description of the tutorial and the material to be distributed to participants. Special session proposals must include a topical title, rationale, session outline, contact information, and a list of invited speakers. Additional information can be found at the ICASSP 2016 website.

Signal Processing Letters: Authors of IEEE Signal Processing Letters (SPL) papers will be given the opportunity to present their work at ICASSP 2016, subject to space availability and approval by the ICASSP Technical Program Chairs. SPL papers published between January 1, 2015 and December 31, 2015 are eligible for presentation at ICASSP 2016.

Show and Tell: S&T offers a perfect stage to showcase innovative ideas in all technical areas of interest at ICASSP. S&T sessions contain demos that are highly interactive and visible. Please refer to the ICASSP 2016 website for additional information regarding demo submission.

Important Deadlines:

Special session & tutorial proposals	August 3, 2015
Notification of special session & tutorial acceptance	September 11, 2015
Submission of regular papers	September 25, 2015
Signal processing letters	December 16, 2015
Notification of paper acceptance	December 21, 2015
Revised paper upload	January 22, 2016
Author registration	January 22, 2016

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

www.ieee.org/sp/index.html

LETTERS

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

www.ieee.org/sp/index.html

LETTERS

2015 IEEE International Workshop on Multimedia Signal Processing

Xiamen, China, October 19 – October 21, 2015
http://www.mmsp2015.org

General Chairs

Xiao-Ping Zhang – *Ryerson U, Canada*

Oscar C. Au – *HKUST, Hong Kong*

Jonathan Li – *Xiamen U, China*

Technical Chairs

Tao Mei – *Microsoft Research Asia*

Gene Cheung – *NII, Japan*

Special Session Chairs

John Paisley – *Columbia U, USA*

Yap-Peng Tan – *NTU, Singapore*

Overview Chairs

Homer Chen–*NTU, Taiwan*

Anthony Vetro – *MERL, USA*

Local Arrangement Chair

Xinghao Ding –*Xiamen U, China*

Rongrong Ji – *Xiamen U, China*

Finance Chairs

Chia-Wen Lin – *NTHU, Taiwan*

Yue Huang – *Xiamen U, China*

Publications Chairs

Vicky Zhao – *U. Alberta, Canada*

Delu Zeng – *Xiamen U, China*

Publicity Chairs

Lina Stankovic – *U. Strathclyde, UK*

Ivan Bajic – *Simon Fraser U,. Canada*

Registration Chair

Liujuan Cao – *Xiamen U, China*

Demo Chair

Wenxin Hong – *Xiamen U, China*

Industry Liaison

Alexander Loui – *Kodak, USA*

North America Liaison

Antonio Ortega, *USC, USA*

Asia Liason

Feng Wu – *USTC, China*

Europe Liaison

Fernando Pereira – *IST-IT, Portugal*

Tentative Call for Papers

MMSP 2015 is the 17th International Workshop on Multimedia Signal Processing. The workshop is organized by the Multimedia Signal Processing Technical Committee of the IEEE Signal Processing Society. This year's event has a ***Heterogeneous Big Data Analytics in Multimedia*** theme. The workshop will bring together researchers and developers in multimedia signal processing and applications to share their latest achievements and explore future directions and synergies in these exciting areas.

Papers are solicited in (but not limited to) the following topics, covering this year's theme and the general scope of multimedia signal processing:

➢ Theories and applications for heterogeneous big media data analytics

➢ Semantic extraction and knowledge mining from heterogeneous big media data

➢ Massive-scale media detection and recognition

➢ Content-based analysis, retrieval and annotation for big media data

➢ Feature learning for heterogeneous big media data representation

➢ Multimedia security, forensic, privacy for big data

➢ Multimedia quality assessment and enhancement

➢ Affective computing and cross-media sentiment analysis

➢ Media algorithm optimization and complexity analysis

➢ Multimedia in economics, finance, business analytics

➢ Multimedia signals in geomatics

➢ Image/video coding and processing

➢ Speech/audio recognition and processing

➢ Multimedia communications and interactions

Top 10% Paper Award

This award is granted to as many as 10% of the total paper submissions, and is open to all accepted papers. Papers will be evaluated based on originality, technical contribution, and presentation quality during the workshop.

Paper Submission

Prospective authors should submit full-length papers of 6 pages in two-column IEEE format, including author affiliation and address, figures, tables and references, to the submission website. Only electronic submissions are accepted. Paper submission implies the intent of at least one of the authors to register and present the paper, if accepted.

Important Dates

Proposals for Special Sessions:	March 20, 2015
Submission of Paper:	May 28, 2015
Notification of acceptance:	July 6, 2015

IEEE SignalProcessing MAGAZINE

SPECTRUM SENSING FOR COGNITIVE RADIO

ACOUSTIC SCENE CLASSIFICATION

DEEP LEARNING IN SPEECH SYNTHESIS

VISUAL DOMAIN ADAPTATION

RESAMPLING METHODS FOR PARTICLE FILTERING

PHASE RETRIEVAL IN OPTICAL IMAGING

PRECIPITATION MONITORING

OPTIMIZATION TOOLS FOR CONFERENCE REVIEW MANAGEMENT

IEEE Signal Processing Society

IEEE

CONTENTS

FREE SPS STUDENT MEMBERSHIP FOR 2015

You're in the beginning stages of your career. Membership in the IEEE Signal Processing Society can help you lay the groundwork for many years of success. You can have it all in 2015 - and for free! Membership includes:

- **Discounts** on conference registration fees;
- Eligibility to apply for **travel grants** to attend SPS flagship conferences including the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) and IEEE International Conference on Image Processing (ICIP);
- **Networking** and **job opportunities** at the ICASSP Student Career Luncheon;
- Eligibility to enter our **student competition**, the Signal Processing Cup, for a US$5,000 grand prize;
- **Involvement opportunities** through SPS's local Chapters - more than 130 worldwide;
- **Free** electronic and digital subscriptions to *IEEE Signal Processing Magazine*, *Inside Signal Processing eNewsletter*, and the *IEEE Signal Processing Society Content Gazette*;
- Access to cutting-edge **educational resources**, including SigView, SPS's online video tutorial portal.

> **See everything Signal Processing Society membership can do for you:**
> http://signalprocessingsociety.org

Already an IEEE member? Join SPS for free now!

(You must have already renewed your IEEE membership for 2015 to use this offer)

- Visit http://ieee.org/join
- On the left side, click "Societies and Special Interest Groups"
- Click "IEEE Signal Processing Society," then "Join the IEEE Signal Processing Society"
- When you reach the catalog page, click "Add Item(s)" and sign in with your IEEE account *Note: Free offer applies only to basic membership. For US$8.00, enhance your membership for more great benefits!*
- Once logged in, click "Proceed to Checkout"
- When you reach the shopping cart, enter the promotion code SP15STUAD and click "Apply"
- Complete check out and congratulations! Welcome to SPS!

Not yet an IEEE Student Member?

Get a free SPS membership with the purchase of an IEEE Student membership!

- Visit http://ieee.org/join
- Click "Join as a student" on the bottom right to create your new IEEE Student member account
- After your IEEE account is created, complete the membership application and proceed to "Do you want to add any memberships and subscriptions?"
- Select "Signal Processing Society membership" and click "add selected item"
- Click "Proceed to Checkout"
- When you reach the shopping cart, enter the promotion code SP15STUNW and click "Apply"
- Complete check out and congratulations! Welcome to SPS!

Note: Must be an active IEEE Student or Graduate Student member. This offer does not apply to SPS Students or Graduate Students renewing for 2015.

◆IEEE

Preliminary Call for Papers

**ISSPIT 2015
Dec., 7-10, 2015
Abu Dhabi
UAE**

The 15th IEEE International Symposium on Signal Processing and Information Technology

December 7-10, 2015, Abu Dhabi, UAE

IEEE   IEEE Signal Processing Society   IEEE Computer Society

Co-sponsored by IEEE Signal Processing Society and IEEE Computer Society

Supported by   Abu Dhabi Education Council   ABU DHABI UNIVERSITY   University of Windsor

General Co-Chairs
Adel Elmaghraby
University of Louisville, USA
Begoña García Zapirain
DeustoTech, Spain

Technical Prog. Chair
Esam Abdel-Raheem
University of Windsor, Canada

Technical Prog. Co-Chairs
Mohammed Ghazal
Abu Dhabi University, UAE
Walaa Sheta
City for Scientific Research, Egypt

Registration & Finance Chair
Reda Ammar
University of Connecticut, USA

Publication Co-Chair s
Hassan Hajjdiab
Abu Dhabi University, UAE
Zakaria Maamar
Zayed University, UAE

Tutorials Co-Chairs
Murad Elhadef
Abu Dhabi University, UAE
Sartaj Sahni
University of Florida, USA

Plenary & Special Sessions Chair
Christos Douligeris
University of Piraeus, Greece

Publicity Co-Chairs
Muhammed Akmal
Abu Dhabi University, UAE
Marco Re
Univ. of Rome "Tor Vergata", Italy

Industrial Liaison Chair
Montasir Qasymeh
Abu Dhabi University, UAE

Local Arrangements Co-Chairs
Riad Kanan
Ashraf Khalil
Abu Dhabi University, UAE

Web Manager
Mostafa G. Mostafa
Mckendree University, USA

The *IEEE ISSPIT 2015* is the fifteenth in a series of international symposia that aims to cover most of the aspects in the fields of signal processing and information technology. Sessions will include tutorials in addition to presentations on new research results. Papers describing original work are invited in any of the areas listed below. Accepted papers will be published in the *Proceedings of IEEE ISSPIT 2015* and will be available via *IEEE Xplore*. Acceptance will be based on quality, relevance, and originality. Contest for Best Paper Award will be held and award will be given.

Papers are invited in the following **topics**:

- Signal Processing Theory and Methods
- Signal Processing for Communications and Networking
- Design & Implementation of Signal Processing Systems
- Image, Video & Multidimensional Signal Processing
- Multimedia Signal Processing
- Biological Image and signal processing
- Audio and Acoustic signal Processing
- Health Informatics and e-Health

- Sensor Arrays
- Radar Signal Processing
- Internet Software Architectures
- Multimedia and Image Based Systems
- Mobile Computing and Applications
- E-Commerce
- Bioinformatics and Bioengineering
- Information Processing
- Geographical Information Systems
- Object Based Software Engineering
- Speech Processing
- Computer Networks
- Neural Networks

Prospective authors are invited to submit full-length, 6-page (max) papers in two-column formats including diagrams and references. Authors can submit their papers as PDF files through the online submission system found on the ISSPIT website: **www.isspit.org**. The title page should include author(s) name(s), affiliation, mailing address, telephone, fax, and e-mail address. The author should indicate one or two of the above categories that best describe the topic of the paper.

Important Dates

Proposals for Tutorials & Special Sessions	Sept. 4th, 2015
Regular paper submission	Sept. 4th, 2015
Notification of acceptance	Oct. 9th, 2015
Final version paper with registration	Oct. 23rd, 2015

website: http://www.isspit.org

IEEE Final Approval Pending

General Chairs

Petar M. Djurić
petar.djuric@stonybrook.edu
Stony Brook University,
USA

Jean-Yves Tourneret
jean-yves.tourneret@enseeiht.fr
University of Toulouse,
France

Technical Program Chairs

Fulvio Gini
f.gini@ing.unipi.it
University of Pisa,
Italy

Cédric Richard
cedric.richard@unice.fr
Nice Sophia-Antipolis University,
France

Finance Chair

Marius Pesavento
mpesa@nt.tu-darmstadt.de
University of Darmstadt,
Germany

Special Sessions Chair

Maria Sabrina Greco
m.greco@iet.unipi.it
University of Pisa,
Italy

Publicity and Publications Chair

Waheed U. Bajwa
waheed.bajwa@rutgers.edu
Rutgers University,
USA

Local Arrangements Chair

Mónica F. Bugallo
monica.bugallo@stonybrook.edu
Stony Brook University,
USA

Join Twitter Conversation
#CAMSAP2015

IEEE

CAMSAP 2015

Call for Papers

The Sixth IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing

Cancun, Mexico
December 13 – 16, 2015
http://inspire.rutgers.edu/CAMSAP2015

Following the success of the first five editions of the IEEE workshop on Computational Advances in Multi-Sensor Adaptive Processing, we are pleased to announce the sixth workshop in this series. IEEE CAMSAP 2015 will be held in historic Cancun, Mexico, and will feature a number of plenary talks from the world's leading researchers in the area, special focus sessions, and contributed papers. All papers will undergo peer review in order to provide feedback to the authors and ensure a high-quality program.

Topics and applications of interest for the workshop include, but are not limited to, the following.

TOPICS OF INTEREST

- Array processing, waveform diversity, space-time adaptive processing
- Convex optimization and relaxation
- Computational linear & multi-linear algebra
- Computer-intensive methods in signal processing (bootstrap, MCMC, EM, particle filtering, etc.)
- Signal and information processing over networks
- Sparse signal processing

APPLICATIONS

- Big data
- Biomedical signal processing
- Communication systems
- Computational imaging
- Radar
- Sensor networks
- Smart grids
- Sonar

Submission of Papers: Prospective authors are invited to submit original full-length papers, with up to four pages for technical content including figures and references, using the formatting guidelines on the website for reviewing purposes. All accepted papers must be presented at the workshop to appear in the proceedings. Best student paper awards, selected by a CAMSAP committee, will also be presented at the workshop.

Special Session Proposals: In addition to contributed sessions, the workshop will also have a number of special sessions. Prospective organizers of special sessions are invited to submit a proposal form, available on the workshop website, by e-mail to the Special Sessions Chair.

IMPORTANT DEADLINES

Submission of proposals for special sessions	February 15, 2015
Notification of special session acceptance	March 15, 2015
Submission of papers	June 15, 2015
Notification of paper acceptance	September 11, 2015
Final paper submission	October 11, 2015

IEEE
Signal Processing Society
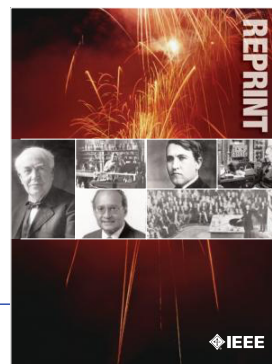
IEEE ORDER FORM FOR REPRINTS

Purchasing IEEE Papers in Print is easy, cost-effective and quick.
Complete this form, send via our secure fax (24 hours a day) to 732-981-8062 or mail it back to us.

PLEASE FILL OUT THE FOLLOWING

Author: _____

Publication Title: _____

Paper Title: _____

_____

RETURN THIS FORM TO:
IEEE Publishing Services
445 Hoes Lane
Piscataway, NJ 08855-1331

Email the Reprint Department at reprints@ieee.org for questions regarding this form

PLEASE SEND ME

■ 50 ■ 100 ■ 200 ■ 300 ■ 400 ■ 500 or _____ (in multiples of 50) reprints.

■ YES ■ NO Self-covering/title page required. COVER PRICE: $74 per 100, $39 per 50.

■ $58.00 Air Freight must be added for all orders being shipped outside the U.S.

■ $21.50 must be added for all USA shipments to cover the cost of UPS shipping and handling.

PAYMENT

■ Check enclosed. Payable on a bank in the USA.

■ Charge my: ■ Visa ■ Mastercard ■ Amex ■ Diners Club

Account # _____ Exp. date _____

Cardholder's Name (please print): _____

■ Bill me (you must attach a purchase order) Purchase Order Number _____

Send Reprints to:

_____

_____

_____

Bill to address, if different:

_____

_____

_____

Because information and papers are gathered from various sources, there may be a delay in receiving your reprint request. This is especially true with postconference publications. Please provide us with contact information if you would like notification of a delay of more than 12 weeks.

Telephone: _____ Fax: _____ Email Address: _____

2012 REPRINT PRICES (without covers)

Number of Text Pages

	1-4	5-8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40	41-44	45-48
50	$129	$213	$245	$248	$288	$340	$371	$408	$440	$477	$510	$543
100	$245	$425	$479	$495	$573	$680	$742	$817	$885	$953	$1021	$1088

Larger quantities can be ordered. Email reprints@ieee.org with specific details.

Tax Applies on shipments of regular reprints to CA, DC, FL, MI, NJ, NY, OH and Canada (GST Registration no. 12534188).
Prices are based on black & white printing. Please call us for full color price quote, if applicable.

2015 IEEE MEMBERSHIP APPLICATION

(students and graduate students must apply online)

IEEE
Advancing Technology
for Humanity

Start your membership immediately: Join online www.ieee.org/join

Please complete both sides of this form, typing or **printing in capital letters**. Use only English characters and abbreviate only if more than 40 characters and spaces per line. We regret that incomplete applications cannot be processed.

1   Name & Contact Information

Please PRINT your name as you want it to appear on your membership card and IEEE correspondence. As a key identifier for the IEEE database, circle your last/surname.

☐ Male   ☐ Female   Date of birth (Day/Month/Year) _____ / _____ / _____

_____
Title   First/Given Name   Middle   Last/Family Surname

▼ Primary Address   ☐ Home   ☐ Business   (All IEEE mail sent here)

_____
Street Address

_____
City   State/Province

_____
Postal Code   Country

_____
Primary Phone

_____
Primary E-mail

▼ Secondary Address   ☐ Home   ☐ Business

_____
Company Name   Department/Division

_____
Street Address   City   State/Province

_____
Postal Code   Country

_____
Secondary Phone

_____
Secondary E-mail

To better serve our members and supplement member dues, your postal mailing address is made available to carefully selected organizations to provide you with information on technical services, continuing education, and conferences. Your e-mail address is not rented by IEEE. Please check box only if you do not want to receive these postal mailings to the selected address. ☐

2   Attestation

I have graduated from a three- to five-year academic program with a university-level degree.
☐ Yes   ☐ No

This program is in one of the following fields of study:
☐ Engineering
☐ Computer Sciences and Information Technologies
☐ Physical Sciences
☐ Biological and Medical Sciences
☐ Mathematics
☐ Technical Communications, Education, Management, Law and Policy
☐ Other (please specify): _____

This academic institution or program is accredited in the country where the institution is located.   ☐ Yes   ☐ No   ☐ Do not know

I have _____ years of professional experience in teaching, creating, developing, practicing, or managing within the following field:
☐ Engineering
☐ Computer Sciences and Information Technologies
☐ Physical Sciences
☐ Biological and Medical Sciences
☐ Mathematics
☐ Technical Communications, Education, Management, Law and Policy
☐ Other (please specify): _____

3   Please Tell Us About Yourself

Select the numbered option that best describes yourself. This information is used by IEEE magazines to verify their annual circulation. Please enter numbered selections in the boxes provided.

A. Primary line of business ⟶ ☐

1. Computers
2. Computer peripheral equipment
3. Software
4. Office and business machines
5. Test, measurement and instrumentation equipment
6. Communications systems and equipment
7. Navigation and guidance systems and equipment
8. Consumer electronics/appliances
9. Industrial equipment, controls and systems
10. ICs and microprocessors
11. Semiconductors, components, sub-assemblies, materials and supplies
12. Aircraft, missiles, space and ground support equipment
13. Oceanography and support equipment
14. Medical electronic equipment
15. OEM incorporating electronics in their end product (not elsewhere classified)
16. Independent and university research, test and design laboratories and consultants (not connected with a mfg. co.)
17. Government agencies and armed forces
18. Companies using and/or incorporating any electronic products in their manufacturing, processing, research or development activities
19. Telecommunications services, telephone (including cellular)
20. Broadcast services (TV, cable, radio)
21. Transportation services (airline, railroad, etc.)
22. Computer and communications and data processing services
23. Power production, generation, transmission and distribution
24. Other commercial users of electrical, electronic equipment and services (not elsewhere classified)
25. Distributor (reseller, wholesaler, retailer)
26. University, college/other educational institutions, libraries
27. Retired
28. Other_____

B. Principal job function ⟶ ☐

1. General and corporate management
2. Engineering management
3. Project engineering management
4. Research and development management
5. Design engineering management —analog
6. Design engineering management —digital
7. Research and development engineering
8. Design/development engineering —analog
9. Design/development engineering—digital
10. Hardware engineering
11. Software design/development
12. Computer science
13. Science/physics/mathematics
14. Engineering (not elsewhere specified)
15. Marketing/sales/purchasing
16. Consulting
17. Education/teaching
18. Retired
19. Other _____

C. Principal responsibility ⟶ ☐

1. Engineering and scientific management
2. Management other than engineering
3. Engineering design
4. Engineering
5. Software: science/mngmnt/engineering
6. Education/teaching
7. Consulting
8. Retired
9. Other _____

D. Title ⟶ ☐

1. Chairman of the Board/President/CEO
2. Owner/Partner
3. General Manager
4. VP Operations
5. VP Engineering/Dir. Engineering
6. Chief Engineer/Chief Scientist
7. Engineering Management
8. Scientific Management
9. Member of Technical Staff
10. Design Engineering Manager
11. Design Engineer
12. Hardware Engineer
13. Software Engineer
14. Computer Scientist
15. Dean/Professor/Instructor
16. Consultant
17. Retired
18. Other _____

Are you now or were you ever a member of IEEE?
☐ Yes   ☐ No   If yes, provide, if known:

_____
Membership Number   Grade   Year Expired

4   Please Sign Your Application

I hereby apply for IEEE membership and agree to be governed by the IEEE Constitution, Bylaws, and Code of Ethics. I understand that IEEE will communicate with me regarding my individual membership and all related benefits. **Application must be signed.**

_____
Signature   Date

Over Please

5   Add IEEE Society Memberships (Optional)

The 39 IEEE Societies support your technical and professional interests. Many society memberships include a personal subscription to the core journal, magazine, or newsletter of that society. **For a complete list of everything included with your IEEE Society membership, visit www.ieee.org/join.** All prices are quoted in US dollars.

Please check ✔ the appropriate box.

Society	Code	BETWEEN 16 AUG 2014-28 FEB 2015 PAY	BETWEEN 1 MAR 2015-15 AUG 2015 PAY
IEEE Aerospace and Electronic Systems ⌘ ⊡	AES010	25.00 ☐	12.50 ☐
IEEE Antennas and Propagation ⌘ ⊡	AP003	15.00 ☐	7.50 ☐
IEEE Broadcast Technology ▪ ⌘ ⊡	BT002	15.00 ☐	7.50 ☐
IEEE Circuits and Systems ▪ ⌘ ⊡	CAS004	19.00 ☐	9.50 ☐
IEEE Communications ⌘ ⊡	COM019	30.00 ☐	15.00 ☐
IEEE Components, Packaging, & Manu. Tech. ▪ ⌘ ⊡	CPMT021	15.00 ☐	7.50 ☐
IEEE Computational Intelligence ⌘ ⊡	CIS011	29.00 ☐	14.50 ☐
IEEE Computer ▪ ⌘ ⊡	C016	56.00 ☐	28.00 ☐
IEEE Consumer Electronics ⌘ ⊡	CE008	20.00 ☐	10.00 ☐
IEEE Control Systems ⌘ ⊡	CS023	25.00 ☐	12.50 ☐
IEEE Dielectrics and Electrical Insulation ⌘ ⊡	DEI032	26.00 ☐	13.00 ☐
IEEE Education ⌘ ⊡	E025	20.00 ☐	10.00 ☐
IEEE Electromagnetic Compatibility ▪ ⌘ ⊡	EMC027	31.00 ☐	15.50 ☐
IEEE Electron Devices ▪ ⌘ ⊡	ED015	18.00 ☐	9.00 ☐
IEEE Engineering in Medicine and Biology ⌘ ⊡	EMB018	40.00 ☐	20.00 ☐
IEEE Geoscience and Remote Sensing ⌘ ⊡	GRS029	19.00 ☐	9.50 ☐
IEEE Industrial Electronics ⌘ ⊡	IE013	15.00 ☐	7.50 ☐
IEEE Industry Applications ⌘ ⊡	IA034	20.00 ☐	10.00 ☐
IEEE Information Theory ▪ ⌘ ⊡	IT012	30.00 ☐	15.00 ☐
IEEE Instrumentation and Measurement ⌘ ⊡	IM009	29.00 ☐	14.50 ☐
IEEE Intelligent Transportation Systems ▪ ⌘ ⊡	ITSS038	35.00 ☐	17.50 ☐
IEEE Magnetics ▪ ⊡ ♦ ⌘	MAG033	26.00 ☐	13.00 ☐
IEEE Microwave Theory and Techniques ⌘ ⊡	MTT017	17.00 ☐	8.50 ☐
IEEE Nuclear and Plasma Sciences ▪ ⌘ ⊡	NPS005	35.00 ☐	17.50 ☐
IEEE Oceanic Engineering ▪ ⌘ ⊡	OE022	19.00 ☐	9.50 ☐
IEEE Photonics ▪ ⌘ ⊡	PHO036	34.00 ☐	17.00 ☐
IEEE Power Electronics ⌘ ⊡	PEL035	25.00 ☐	12.50 ☐
IEEE Power & Energy ▪ ⌘ ⊡	PE031	35.00 ☐	17.50 ☐
IEEE Product Safety Engineering ▪	PSE043	35.00 ☐	17.50 ☐
IEEE Professional Communication ▪ ⌘ ⊡	PC026	31.00 ☐	15.50 ☐
IEEE Reliability ▪ ⌘ ⊡	RL007	35.00 ☐	17.50 ☐
IEEE Robotics and Automation ⌘ ⊡	RA024	9.00 ☐	4.50 ☐
IEEE Signal Processing ▪ ⌘ ⊡	SP001	20.00 ☐	10.00 ☐
IEEE Social Implications of Technology ▪ ⌘ ⊡	SIT030	33.00 ☐	16.50 ☐
IEEE Solid-State Circuits ⌘ ⊡	SSC037	29.00 ☐	14.50 ☐
IEEE Systems, Man, & Cybernetics ⌘ ⊡	SMC028	12.00 ☐	6.00 ☐
IEEE Technology & Engineering Management ⌘ ⊡	TEM014	35.00 ☐	17.50 ☐
IEEE Ultrasonics, Ferroelectrics, & Frequency Control ▪ ⌘ ⊡	UFFC020	20.00 ☐	10.00 ☐
IEEE Vehicular Technology ▪ ⌘ ⊡	VT006	18.00 ☐	9.00 ☐

Legend—Society membership includes:

⌘ One or more Society publications   ⊡ Online access to publication

▪ Society newsletter   ♦ CD-ROM of selected society publications

Complete both sides of this form, sign, and return to:
IEEE MEMBERSHIP APPLICATION PROCESSING
445 HOES LN, PISCATAWAY, NJ 08854-4141 USA
or fax to +1 732 981 0225
or join online at www.ieee.org/join

Please reprint your full name here

6   2015 IEEE Membership Rates
(student rates available online)

IEEE member dues and regional assessments are based on where you live and when you apply. Membership is based on the calendar year from 1 January through 31 December. All prices are quoted in US dollars.

Please check ✔ the appropriate box.

RESIDENCE	BETWEEN 16 AUG 2014-28 FEB 2015 PAY	BETWEEN 1 MAR 2015-15 AUG 2015 PAY
United States	$193.00 ☐	$96.50 ☐
Canada (GST)*	$171.25 ☐	$85.63 ☐
Canada (NB, NF and ON HST)*	$182.85 ☐	$91.43 ☐
Canada (Nova Scotia HST)*	$185.75 ☐	$92.88 ☐
Canada (PEI HST)*	$184.30 ☐	$92.15 ☐
Canada (GST and QST Quebec)	$185.71 ☐	$92.86 ☐
Africa, Europe, Middle East	$158.00 ☐	$79.00 ☐
Latin America	$149.00 ☐	$74.50 ☐
Asia, Pacific	$150.00 ☐	$75.00 ☐

*IEEE Canada Business No. 125634188

Minimum Income or Unemployed Provision
Applicants who certify that their prior year income did not exceed US$14,500 (or equivalent) or were not employed are granted 50% reduction in: full-year dues, regional assessment and fees for one IEEE Membership plus one Society Membership. If applicable, please check appropriate box and adjust payment accordingly. Student members are not eligible.

☐ I certify I earned less than US$14,500 in 2014 or 2013
☐ I certify that I was unemployed in 2014 or 2013

7   More Recommended Options

Proceedings of the IEEE .... print $45.00 ☐ **or** online $39.00 ☐
Proceedings of the IEEE (print/online combination) ....$55.00 ☐
IEEE Standards Association (IEEE-SA) ....$52.00 ☐
IEEE Women in Engineering (WIE) ....$25.00 ☐

8   Payment Amount

Please total the Membership dues, Society dues, and other amounts from this page:

IEEE Membership dues   ⑥ ....$_____
IEEE Society dues (optional)   ⑤ ....$_____
IEEE-SA/WIE dues (optional)   ⑦ ....$_____
Proceedings of the IEEE (optional)   ⑦ ....$_____
Canadian residents pay 5% GST or appropriate HST (BC—12%; NB, NF, ON-13%;NS-15%) on Society payments & publications only....................TAX $_____
AMOUNT PAID .... **TOTAL** $_____

Payment Method
All prices are quoted in US dollars. You may pay for IEEE membership by credit card (see below), check, or money order payable to IEEE, drawn on a US bank.

☐ Check   ☐ VISA   ☐ MasterCard   ☐ Diners Club International   ☐ American Express   ☐ Discover

Credit Card Number

MONTH   YEAR
EXPIRATION DATE   CARDHOLDER'S 5-DIGIT ZIPCODE (BILLING STATEMENT ADDRESS) USA ONLY

Name as it appears on card

Signature

Auto Renew my Memberships and Subscriptions (available when paying by credit card).
☐ I agree to the Terms and Conditions located at www.ieee.org/autorenew

9   Were You Referred to IEEE?

☐ Yes   ☐ No   If yes, provide the following:
Member Recruiter Name _____
IEEE Recruiter's Member Number (Required) _____

CAMPAIGN CODE   PROMO CODE

7/14

14-MEM-119

Information for Authors

(Updated/Effective January 2015)

For Transactions and Journals:

Authors are encouraged to submit manuscripts of Regular papers (papers which provide a complete disclosure of a technical premise), or Comment Correspondences (brief items that provide comment on a paper previously published in these TRANSACTIONS). Submissions/resubmissions must be previously unpublished and may not be under consideration elsewhere.

Every manuscript must:

i. provide a clear statement of the problem and what the contribution of the work is to the relevant research community;
ii. state why this contribution is significant (what impact it will have);
iii. provide citation of the published literature most closely related to the manuscript; and
iv. state what is distinctive and new about the current manuscript relative to these previously published works.

By submission of your manuscript to these TRANSACTIONS, all listed authors have agreed to the authorship list and all the contents and confirm that the work is original and that figures, tables and other reported results accurately reflect the experimental work. In addition, the authors all acknowledge that they accept the rules established for publication of manuscripts, including agreement to pay all overlength page charges, color charges, and any other charges and fees associated with publication of the manuscript. Such charges are not negotiable and cannot be suspended. The corresponding author is responsible for obtaining consent from all co-authors and, if needed, from sponsors before submission.

In order to be considered for review, a paper must be within the scope of the journal and represent a novel contribution. A paper is a candidate for an Immediate Rejection if it is of limited novelty, e.g. a straightforward combination of theories and algorithms that are well established and are repeated on a known scenario. Experimental contributions will be rejected without review if there is insufficient experimental data. These TRANSACTIONS are published in English. Papers that have a large number of typographical and/or grammatical errors will also be rejected without review.

In addition to presenting a novel contribution, acceptable manuscripts must describe and cite related work in the field to put the contribution in context. Do not give theoretical derivations or algorithm descriptions that are easily found in the literature; merely cite the reference.

New and revised manuscripts should be prepared following the "Manuscript Submission" guidelines below, and submitted to the online manuscript system, ScholarOne Manuscripts. Do not send original submissions or revisions directly to the Editor-in-Chief or Associate Editors; they will access your manuscript electronically via the ScholarOne Manuscript system.

Manuscript Submission. Please follow the next steps.

1. *Account in ScholarOne Manuscripts.* If necessary, create an account in the on-line submission system ScholarOne Manuscripts. Please check first if you already have an existing account which is based on your e-mail address and may have been created for you when you reviewed or authored a previous paper.
2. *Electronic Manuscript.* Prepare a PDF file containing your manuscript in double-column, single-spaced format using a font size of 10 points or larger, having a margin of at least 1 inch on all sides. Upload this version of the manuscript as a PDF file "double.pdf" to the ScholarOne-Manuscripts site. Since many reviewers prefer a larger font, you are strongly encouraged to also submit a single-column, double-spaced version (11 point font or larger), which is easy to create with the templates provided **IEEE Author Digital Toolbox (http://www.ieee.org/publications_standards/publications/authors/authors_journals.html)**. Page length restrictions will be determined by the double-column

_____
Digital Object Identifier

version. Proofread your submission, confirming that all figures and equations are visible in your document before you "SUBMIT" your manuscript. Proofreading is critical; once you submit your manuscript, the manuscript cannot be changed in any way. You may also submit your manuscript as a .PDF or MS Word file. The system has the capability of converting your files to PDF, however it is your responsibility to confirm that the conversion is correct and there are no font or graphics issues prior to completing the submission process.

3. *EDICS (Not applicable to Journal of Selected Topics in Signal Processing).* All submissions must be classified by the author with an EDICS (Editors' Information Classification Scheme) selected from the list of EDICS published online at the at the publication's EDICS webpage (*please see the list below). Upon submission of a new manuscript, please choose the EDICS categories that best suit your manuscript. Failure to do so will likely result in a delay of the peer review process.

4. *Additional Documents for Review.* Please upload pdf versions of all items in the reference list that are not publicly available, such as unpublished (submitted) papers. Graphical abstracts and supplemental materials intended to appear with the final paper (see below) must also be uploaded for review at the time of the initial submission for consideration in the review process. Use short filenames without spaces or special characters. When the upload of each file is completed, you will be asked to provide a description of that file.

5. *Supplemental Materials.* IEEE Xplore can publish multimedia files (audio, images, video), datasets, and software (e.g. Matlab code) along with your paper. Alternatively, you can provide the links to such files in a README file that appears on Xplore along with your paper. For details, please see IEEE Author Digital Toolbox under "Multimedia." To make your work reproducible by others, these TRANSACTIONS encourages you to submit all files that can recreate the figures in your paper.

6. *Submission.* After uploading all files and proofreading them, submit your manuscript by clicking "Submit." A confirmation of the successful submission will open on screen containing the manuscript tracking number and will be followed with an e-mail confirmation to the corresponding and all contributing authors. Once you click "Submit," your manuscript cannot be changed in any way.

7. *Copyright Form and Consent Form.* By policy, IEEE owns the copyright to the technical contributions it publishes on behalf of the interests of the IEEE, its authors, and their employers; and to facilitate the appropriate reuse of this material by others. To comply with the IEEE copyright policies, authors are required to sign and submit a completed "IEEE Copyright and Consent Form" prior to publication by the IEEE. The IEEE recommends authors to use an effective electronic copyright form (eCF) tool within the ScholarOne Manuscripts system. You will be redirected to the "IEEE Electronic Copyright Form" wizard at the end of your original submission; please simply sign the eCF by typing your name at the proper location and click on the "Submit" button.

Comment Correspondence. Comment Correspondences provide brief comments on material previously published in these TRANSACTIONS. These items may not exceed 2 pages in double-column, single spaced format, using 9 point type, with margins of 1 inch minimum on all sides, and including: title, names and contact information for authors, abstract, text, references, and an appropriate number of illustrations and/or tables. Correspondence items are submitted in the same way as regular manuscripts (see "Manuscript Submission" above for instructions).

Authors may also submit manuscripts of overview articles, but note that these include an additional white paper approval process http://www.signalprocessingsociety.org/publications/overview-articles/. *[This does not apply to the Journal of Selected Topics in Signal Processing. Please contact the Editor-in-Chief.]*

Manuscript Length. For the initial submission of a regular paper, the manuscript may not exceed 13 double-column pages (10 point font), including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs; and all references. Supplemental materials and graphical abstracts are not included in the page count. For regular papers, the revised manuscript may not exceed 16 double-column pages (10 point font), including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs; and all references. For Overview Papers, the maximum length is double that for regular submissions at each stage (please reference http://www.signalprocessingsociety.org/publications/overview-articles/ for more information).

Note that any paper in excess of 10 pages will be subject to mandatory overlength page charges. Since changes recommended as a result of peer review may require additions to the manuscript, it is strongly recommended that you practice economy in preparing original submissions. Note: Papers submitted to the TRANSACTIONS ON MULTIMEDIA in excess of 8 pages will be subject to mandatory overlength page charges.

Exceptions to manuscript length requirements may, under extraordinary circumstances, be granted by the Editor-in-Chief. However, such exception does not obviate your requirement to pay any and all overlength or additional charges that attach to the manuscript.

Resubmission of Previously Rejected Manuscripts. Authors of manuscripts rejected from any journal are allowed to resubmit their manuscripts only once. At the time of submission, you will be asked whether your manuscript is a new submission or a resubmission of an earlier rejected manuscript. If it is a resubmission of a manuscript previously rejected by any journal, you are expected to submit supporting documents identifying the previous submission and detailing how your new version addresses all of the reviewers' comments. Papers that do not disclose connection to a previously rejected paper or that do not provide documentation as to changes made may be immediately rejected.

Author Misconduct. Author misconduct includes plagiarism, self-plagiarism, and research misconduct, including falsification or misrepresentation of results. All forms of misconduct are unacceptable and may result in sanctions and/or other corrective actions. Plagiarism includes copying someone else's work without appropriate credit, using someone else's work without clear delineation of citation, and the uncited reuse of an author's previously published work that also involves other authors. Self-plagiarism involves the verbatim copying or reuse of an authors own prior work without appropriate citation, including duplicate submission of a single journal manuscript to two different journals, and submission of two different journal manuscripts which overlap substantially in language or technical contribution. For more information on the definitions, investigation process, and corrective actions related to author misconduct, see the Signal Processing Society Policies and Procedures Manual, Section 6.1. http://www.signalprocessingsociety.org/about-sps/governance/policy-procedure/part-2. Author misconduct may also be actionable by the IEEE under the rules of Member Conduct.

Extensions of the Author's Prior Work. It is acceptable for conference papers to be used as the basis for a more fully developed journal submission. Still, authors are required to cite their related prior work; the papers cannot be identical; and the journal publication must include substantively novel aspects such as new experimental results and analysis or added theoretical work. The journal publication should clearly specify how the journal paper offers novel contributions when citing the prior work. Limited overlap with prior journal publications with a common author is allowed only if it is necessary for the readability of the paper, and the prior work must be cited as the primary source.

Submission Format. Authors are required to prepare manuscripts employing the on-line style files developed by IEEE, which include guidelines for abbreviations, mathematics, and graphics. All manuscripts accepted for publication will require the authors to make final submission employing these style files. The style files are available on the web at the **IEEE Author Digital Toolbox** under "Template for all TRANSACTIONS." (LaTeX and MS Word). Please note the following requirements about the abstract:

- The abstract must be a concise yet comprehensive reflection of what is in your article.
- The abstract must be self-contained, without abbreviations, footnotes, displayed equations, or references.

- The abstract must be between 150-250 words.
- The abstract should include a few keywords or phrases, as this will help readers to find it. Avoid over-repetition of such phrases as this can result in a page being rejected by search engines.

In addition to written abstracts, papers may include a graphical abstract; see http://www.ieee.org/publications_standards/publications/authors/authors_journals.html. for options and format requirements.

IEEE supports the publication of author names in the native language alongside the English versions of the names in the author list of an article. For more information, see "Author names in native languages" (http://www.ieee.org/publications_standards/publications/authors/auth_names_native_lang.pdf) on the IEEE Author Digital Toolbox page.

Open Access. The publication is a hybrid journal, allowing either Traditional manuscript submission or Open Access (author-pays OA) manuscript submission. Upon submission, if you choose to have your manuscript be an Open Access article, you commit to pay the discounted $1,750 OA fee if your manuscript is accepted for publication in order to enable unrestricted public access. Any other application charges (such as overlength page charge and/or charge for the use of color in the print format) will be billed separately once the manuscript formatting is complete but prior to the publication. If you would like your manuscript to be a Traditional submission, your article will be available to qualified subscribers and purchasers via IEEE Xplore. No OA payment is required for Traditional submission.

Page Charges.

Voluntary Page Charges. Upon acceptance of a manuscript for publication, the author(s) or his/her/their company or institution will be asked to pay a charge of $110 per page to cover part of the cost of publication of the first ten pages that comprise the standard length (two pages, in the case of Correspondences).

Mandatory Page Charges The author(s) or his/her/their company or institution will be billed $220 per each page in excess of the first ten published pages for regular papers and six published pages for correspondence items. (**NOTE: Papers accepted to IEEE TRANSACTIONS ON MULTIMEDIA in excess of 8 pages will be subject to mandatory overlength page charges.) These are mandatory page charges and the author(s) will be held responsible for them. They are not negotiable or voluntary. The author(s) signifies his willingness to pay these charges simply by submitting his/her/their manuscript to the TRANSACTIONS. The Publisher holds the right to withhold publication under any circumstance, as well as publication of the current or future submissions of authors who have outstanding mandatory page charge debt. No mandatory overlength page charges will be applied to overview articles in the Society's journals.

Color Charges. Color figures which appear in color only in the electronic (Xplore) version can be used free of charge. In this case, the figure will be printed in the hardcopy version in grayscale, and the author is responsible that the corresponding grayscale figure is intelligible. Color reproduction charges for print are the responsibility of the author. Details of the associated charges can be found on the IEEE Publications page.

Payment of fees on color reproduction is not negotiable or voluntary, and the author's agreement to publish the manuscript in these TRANSACTIONS is considered acceptance of this requirement.

*EDICS Webpages:

IEEE TRANSACTIONS ON SIGNAL PROCESSING:
http://www.signalprocessingsociety.org/publications/periodicals/tsp/TSP-EDICS/
IEEE TRANSACTIONS ON IMAGE PROCESSING:
http://www.signalprocessingsociety.org/publications/periodicals/image-processing/tip-edics/
IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE / ACM:
http://www.signalprocessingsociety.org/publications/periodicals/taslp/taslp-edics/
IEEE TRANSACTIONS ON INFORMATION, FORENSICS AND SECURITY:
http://www.signalprocessingsociety.org/publications/periodicals/forensics/forensics-edics/
IEEE TRANSACTIONS ON MULTIMEDIA:
http://www.signalprocessingsociety.org/tmm/tmm-edics/
IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING:
http://www.signalprocessingsociety.org/publications/periodicals/tci/tci-edics/
IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS:
http://www.signalprocessingsociety.org/publications/periodicals/tsipn/tsipn-edics/

2015 IEEE SIGNAL PROCESSING SOCIETY MEMBERSHIP APPLICATION

Mail to: IEEE OPERATIONS CENTER, ATTN: Louis Curcio, Member and Geographic Activities, 445 Hoes Lane, Piscataway, New Jersey 08854 USA
or Fax to (732) 981-0225 (credit card payments only.)
For info call (732) 981-0060 or 1 (800) 678-IEEE or E-mail: new.membership@ieee.org

◆IEEE

1. PERSONAL INFORMATION

NAME AS IT SHOULD APPEAR ON IEEE MAILINGS: SEND MAIL TO: ☐ Home Address OR ☐ Business/School Address
If not indicated, mail will be sent to home address. Note: Enter your name as you wish it to appear on membership card and all correspondence.
PLEASE PRINT Do not exceed 40 characters or spaces per line. Abbreviate as needed. Please circle your last/surname as a key identifier for the IEEE database.

TITLE   FIRST OR GIVEN NAME   MIDDLE NAME   SURNAME/LAST NAME

HOME ADDRESS

CITY   STATE/PROVINCE   POSTAL CODE   COUNTRY

2.
Are you now or were you ever a member of IEEE? ☐ Yes ☐ No
If yes, please provide, if known:

MEMBERSHIP NUMBER_____

Grade_____Year Membership Expired:_____

3. BUSINESS/PROFESSIONAL INFORMATION

Company Name

Department/Division

Title/Position   Years in Current Position

Years in the Profession Since Graduation   ☐ PE State/Province

Street Address

City   State/Province   Postal Code   Country

4. EDUCATION
A baccalaureate degree from an IEEE recognized educational program assures assignment of "Member" grade. For others, additional information and references may be necessary for grade assignment.

A. _____
Baccalaureate Degree Received   Program/Course of Study

College/University   Campus

State/Province   Country   Mo./Yr. Degree Received

B. _____
Highest Technical Degree Received   Program/Course of Study

College/University   Campus

State/Province   Country   Mo./Yr. Degree Received

5. _____
Full signature of applicant

6. DEMOGRAPHIC INFORMATION   – ALL APPLICANTS -

Date Of Birth _____   ☐ Male ☐ Female
          Day   Month   Year

7. CONTACT INFORMATION

Office Phone/Office Fax   Home Phone/Home Fax

Office E-Mail   Home E-Mail

8. 2015 IEEE MEMBER RATES

IEEE DUES Residence	16 Aug 14-28 Feb 15 Pay Full Year	1 Mar -15 Aug 15 Pay Half Year**
United States	$193.00 ☐	$96.50 ☐
Canada (incl. GST)	$171.25 ☐	$85.63 ☐
Canada (incl. HST for PEI)	$184.30 ☐	$92.15 ☐
Canada (incl. HST for Nova Scotia)	$185.75 ☐	$92.88 ☐
Canada (incl. HST for NB, NF and ON)	$182.85 ☐	$91.43 ☐
Canada (incl. GST and QST Quebec)	$185.71 ☐	$92.86 ☐
Africa, Europe, Middle East	$158.00 ☐	$79.00 ☐
Latin America	$149.00 ☐	$74.50 ☐
Asia, Pacific	$150.00 ☐	$75.00 ☐

Canadian Taxes (GST/HST): All supplies, which include dues, Society membership fees, online products and publications (except CD-ROM and DVD media), shipped to locations within Canada are subject to the GST of 5% or the HST of 13%, 14% or 15%, depending on the Province to which the materials are shipped. GST and HST do not apply to Regional Assessments. (IEEE Canadian Business Number 12563 4188 RT0001)
Value Added Tax (VAT) in the European Union: In accordance with the European Union Council Directives 2002/38/EC and 77/388/EEC amended by Council Regulation (EC)792/2002, IEEE is required to charge and collect VAT on electronic/digitized products sold to private consumers that reside in the European Union. The VAT rate applied is the EU member country standard rate where the consumer is resident. (IEEE's VAT registration number is EU826000081)
U.S. Sales Taxes: Please add applicable state and local sales and use tax on orders shipped to **Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Illinois, Indiana, Kentucky, Massachusetts, Maryland, Michigan, Minnesota, Missouri, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, West Virginia, Wisconsin.** Customers claiming a tax exemption must include an appropriate and properly completed tax-exemption certificate with their first order.

2015 SPS MEMBER RATES

	16 Aug-28 Feb Pay Full Year	1 Mar-15 Aug Pay Half Year
Signal Processing Society Membership Fee*	$ 20.00 ☐	$ 10.00 ☐

Fee includes: IEEE Signal Processing Magazine (electronic and digital), Inside Signal Proc. eNewsletter (electronic) and IEEE Signal Processing Society Content Gazette (electronic).

Add $15 to enhance SPS Membership and also receive:	$ 15.00 ☐	$ 7.50 ☐

IEEE Signal Processing Magazine (print) and SPS Digital Library: online access to Signal Processing Magazine, Signal Processing Letters, Journal of Selected Topics in Signal Processing, Trans. on Audio, Speech, and Language Processing, Trans. on Image Processing, Trans. on Information Forensics and Security and Trans. on Signal Processing.
Publications available only with SPS membership:

Publication			
Signal Processing, IEEE Transactions on:	Print	$190.00 ☐	$ 95.00 ☐
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print	$145.00 ☐	$ 72.50 ☐
Image Processing, IEEE Transactions on:	Print	$188.00 ☐	$ 94.00 ☐
Information Forensics and Security, IEEE Trans. on:	Print	$163.00 ☐	$ 81.50 ☐
IEEE Journal of Selected Topics in Signal Processing:	Print	$160.00 ☐	$ 80.00 ☐
Affective Computing, IEEE Transactions on:	Electronic	$ 35.00 ☐	$ 17.50 ☐
Biomedical and Health Informatics, IEEE Journal of:	Print	$ 55.00 ☐	$ 27.50 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$ 65.00 ☐	$ 32.50 ☐
IEEE Cloud Computing	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
New! IEEE Trans. on Cognitive Comm. &Networking	Electronic	$ 26.00 ☐	$ 13.00 ☐
New! IEEE Trans. on Computational Imaging	Electronic	$ 28.00 ☐	$ 14.00 ☐
New! IEEE Trans. on Big Data	Electronic	$ 25.00 ☐	$ 12.50 ☐
New! IEEE Trans. on Molecular, Biological, & Multi-scale Communications	Electronic	$ 24.00 ☐	$ 12.00 ☐
IEEE Internet of Things Journal	Electronic	$ 26.00 ☐	$ 13.00 ☐
IEEE Trans. on Cloud Computing	Electronic	$ 42.00 ☐	$ 21.00 ☐
IEEE Trans. on Computational Social Systems	Electronic	$ 30.00 ☐	$ 15.00 ☐
New! IEEE Trans. on Signal & Info Proc. Over Networks	Electronic	$ 28.00 ☐	$ 14.00 ☐
IEEE Biometrics Compendium:	Online	$ 30.00 ☐	$ 15.00 ☐
Computing in Science & Engrg. Mag.:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$ 149.00 ☐	$ 74.50 ☐
Medical Imaging, IEEE Transactions on:	Print	$ 74.00 ☐	$ 37.00 ☐
	Electronic	$ 53.00 ☐	$ 26.50 ☐
	Print & Electronic	$ 89.00 ☐	$ 44.50 ☐
Mobile Computing, IEEE Transactions on:	ELE/Print Abstract/CD-ROM	$ 40.00 ☐	$ 20.00 ☐
Multimedia, IEEE Transactions on:	Electronic	$ 42.00 ☐	$ 21.00 ☐
IEEE MultiMedia Magazine:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$149.00 ☐	$ 74.50 ☐
Network Science and Engrg., IEEE Trans. on:	Electronic	$ 33.00 ☐	$ 16.50 ☐
IEEE Reviews in Biomedical Engineering:	Print	$ 25.00 ☐	$ 12.50 ☐
	Electronic	$ 25.00 ☐	$ 12.50 ☐
	Print & Electronic	$ 40.00 ☐	$ 20.00 ☐
IEEE Security and Privacy Magazine:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$149.00 ☐	$ 74.50 ☐
IEEE Sensors Journal:	Print	$150.00 ☐	$ 75.00 ☐
	Electronic	$ 50.00 ☐	$ 25.00 ☐
Smart Grid, IEEE Transactions on:	Print	$100.00 ☐	$ 50.00 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$120.00 ☐	$ 60.00 ☐
Wireless Communications, IEEE Transactions on:	Print	$120.00 ☐	$ 60.00 ☐
	Electronic	$ 48.00 ☐	$ 24.00 ☐
	Print & Electronic	$120.00 ☐	$ 60.00 ☐
IEEE Wireless Communications Letters:	Print	$ 80.00 ☐	$ 40.00 ☐
	Electronic	$ 18.00 ☐	$ 9.00 ☐
	Print & Electronic	$ 95.00 ☐	$ 47.50 ☐
New! IEEE Life Sciences Letters (Open Access Pub)	Electronic	No Fee	

*IEEE membership required or requested
Affiliate application to join SP Society only.   Amount Paid $_____

9.

IEEE Membership Affiliate Fee (See pricing in Section 8)	$_____
Signal Processing Society Fees	$_____

Canadian residents pay 5% GST or 13% HST
Reg. No. 125634188 on Society payment(s) & pubs only   Tax $_____

AMOUNT PAID WITH APPLICATION   TOTAL $_____
Prices subject to change without notice.
☐ Check or money order enclosed Payable to IEEE on a U.S. Bank
☐ American Express   ☐ VISA   ☐ MasterCard
☐ Diners Club

Exp. Date/ Mo./Yr.

Cardholder Zip Code Billing Statement Address/USA Only

Full signature of applicant using credit card   Date

10. WERE YOU REFERRED?
☐ Yes ☐ No   If yes, please provide the follow information:
Member Recruiter Name:_____
IEEE Recruiter's Member Number (Required): _____

IEEE Signal Processing Society

2015 IEEE SIGNAL PROCESSING SOCIETY STUDENT MEMBERSHIP APPLICATION

(Current and reinstating IEEE members joining SPS complete areas 1, 2, 8, 9.)
Mail to: IEEE OPERATIONS CENTER, ATTN: Louis Curcio, Member and Geographic Activities, 445 Hoes Lane, Piscataway, New Jersey 08854 USA
or Fax to (732) 981-0225 (credit card payments only.)
For info call (732) 981-0060 or 1 (800) 678-IEEE or E-mail: new.membership@ieee.org

1. PERSONAL INFORMATION

NAME AS IT SHOULD APPEAR ON IEEE MAILINGS: **SEND MAIL TO:** ☐ Home Address OR ☐ Business/School Address
If not indicated, mail will be sent to home address. Note: Enter your name as you wish it to appear on membership card and all correspondence.
PLEASE PRINT Do not exceed 40 characters or spaces per line. Abbreviate as needed. Please circle your last/surname as a key identifier for the IEEE database.

TITLE   FIRST OR GIVEN NAME   MIDDLE NAME   SURNAME/LAST NAME

HOME ADDRESS

CITY   STATE/PROVINCE   POSTAL CODE   COUNTRY

2. Are you now or were you ever a member of IEEE? ☐ Yes ☐ No
If yes, please provide, if known:

MEMBERSHIP NUMBER_____

Grade_____Year Membership Expired:_____

3. BUSINESS/PROFESSIONAL INFORMATION

Company Name

Department/Division

Title/Position   Years in Current Position

Years in the Profession Since Graduation   ☐ PE State/Province

Street Address

City   State/Province   Postal Code   Country

4. EDUCATION
A baccalaureate degree from an IEEE recognized educational program assures assignment of "Member" grade. For others, additional information and references may be necessary for grade assignment.

A.
Baccalaureate Degree Received   Program/Course of Study

College/University   Campus

State/Province   Country   Mo./Yr. Degree Received

B.
Highest Technical Degree Received   Program/Course of Study

College/University   Campus

State/Province   Country   Mo./Yr. Degree Received

5.
Full signature of applicant

6. DEMOGRAPHIC INFORMATION   – ALL APPLICANTS -
Date Of Birth _____   ☐ Male ☐ Female
Day Month Year

7. CONTACT INFORMATION
Office Phone/Office Fax   Home Phone/Home Fax

Office E-Mail   Home E-Mail

8. 2015 IEEE STUDENT MEMBER RATES

IEEE DUES Residence	16 Aug 14-28 Feb 15 Pay Full Year	1 Mar -15 Aug 15 Pay Half Year**
United States	$32.00 ☐	$16.00 ☐
Canada (incl. GST)	$33.60 ☐	$16.80 ☐
Canada (incl. HST for NB, NF, and ON)	$36.16 ☐	$18.08 ☐
Canada (incl. HST for Nova Scotia)	$36.80 ☐	$18.40 ☐
Canada (incl. HST for PEI)	$36.48 ☐	$18.24 ☐
Canada (incl. GST and QST Quebec)	$36.79 ☐	$18.40 ☐
Africa, Europe, Middle East, Latin America, Asia, Pacific	$27.00 ☐	$13.50 ☐

Canadian Taxes (GST/HST): All supplies, which include dues, Society membership fees, online products and publications (except CD-ROM and DVD media), shipped to locations within Canada are subject to the GST of 5% or the HST of 13%,14% or 15%, depending on the Province to which the materials are shipped. GST and HST do not apply to Regional Assessments. (IEEE Canadian Business Number 12563 4188 RT0001)
Value Added Tax (VAT) in the European Union: In accordance with the European Union Council Directives 2002/38/EC and 77/388/EEC amended by Council Regulation (EC)792/2002, IEEE is required to charge and collect VAT on electronic/digitized products sold to private consumers that reside in the European Union. The VAT rate applied is the EU member country standard rate where the consumer is resident. (IEEE's VAT registration number is EU826000081)
U.S. Sales Taxes: Please add applicable state and local sales and use tax on orders shipped to **Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Illinois, Indiana, Kentucky, Massachusetts, Maryland, Michigan, Minnesota, Missouri, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, West Virginia, Wisconsin.** Customers claiming a tax exemption must include an appropriate and properly completed tax-exemption certificate with their first order.

2015 SPS STUDENT MEMBER RATES

	16 Aug-28 Feb Pay Full Year	1 Mar-15 Aug Pay Half Year
Signal Processing Society Membership Fee*	$10.00 ☐	$ 5.00 ☐

Fee includes: **IEEE Signal Processing Magazine** (electronic and digital), **Inside Signal Processing eNewsletter** (electronic) and **IEEE Signal Processing Society Content Gazette** (electronic).

Add $8 to enhance SPS Membership and also receive: $ 8.00 ☐ $ 4.00 ☐
IEEE Signal Processing Society Magazine (print) and **SPS Digital Library**: online access to Signal Processing Magazine, Signal Processing Letters, Journal of Selected Topics in Signal Processing, Trans. on Audio, Speech, and Language Processing, Trans. on Image Processing, Trans. on Information Forensics and Security and Trans. on Signal Processing.

Publications available only with SPS membership:

		Full Yr	Half Yr
Signal Processing, IEEE Transactions on:	Print	$ 95.00 ☐	$ 47.50 ☐
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print	$ 73.00 ☐	$ 36.50 ☐
Image Processing, IEEE Transactions on:	Print	$ 94.00 ☐	$ 47.00 ☐
Information Forensics and Security, IEEE Trans. on:	Print	$ 82.00 ☐	$ 41.00 ☐
IEEE Journal of Selected Topics in Signal Processing:	Print	$ 80.00 ☐	$ 40.00 ☐
Affective Computing, IEEE Transactions on:	Electronic	$ 18.00 ☐	$ 9.00 ☐
Biomedical and Health Informatics, IEEE Journal of:	Print	$ 28.00 ☐	$ 14.00 ☐
	Electronic	$ 20.00 ☐	$ 10.00 ☐
	Print & Electronic	$ 65.00 ☐	$ 32.50 ☐
IEEE Cloud Computing	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
New! IEEE Trans. on Cognitive Comm. &Networking	Electronic	$ 13.00 ☐	$ 6.50 ☐
New! IEEE Trans. on Computational Imaging	Electronic	$ 14.00 ☐	$ 7.00 ☐
New! IEEE Trans. on Big Data	Electronic	$ 13.00 ☐	$ 6.50 ☐
New! IEEE Trans. on Molecular, Biological, & Multi-Scale Communications	Electronic	$ 12.00 ☐	$ 6.00 ☐
IEEE Internet of Things Journal	Electronic	$ 13.00 ☐	$ 6.50 ☐
IEEE Trans. on Cloud Computing	Electronic	$ 21.00 ☐	$ 10.50 ☐
IEEE Trans. on Computational Social Systems	Electronic	$ 15.00 ☐	$ 7.50 ☐
New! IEEE Trans. on Signal & Info Proc. Over Networks	Electronic	$ 14.00 ☐	$ 7.00 ☐
IEEE Biometrics Compendium:	Online	$ 15.00 ☐	$ 7.50 ☐
Computing in Science & Engrg. Mag.:	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
	Print	$ 75.00 ☐	$ 37.50 ☐
Medical Imaging, IEEE Transactions on:	Print	$ 37.00 ☐	$ 18.50 ☐
	Electronic	$ 27.00 ☐	$ 13.50 ☐
	Print & Electronic	$ 45.00 ☐	$ 22.50 ☐
Mobile Computing, IEEE Transactions on:	ELE/Print Abstract/CD-ROM	$ 20.00 ☐	$ 10.00 ☐
Multimedia, IEEE Transactions on:	Electronic	$ 21.00 ☐	$ 10.50 ☐
IEEE MultiMedia Magazine:	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
	Print	$ 75.00 ☐	$ 37.50 ☐
Network Science and Engrg., IEEE Trans. on:	Electronic	$ 17.00 ☐	$ 8.50 ☐
IEEE Reviews in Biomedical Engineering:	Print	$ 13.00 ☐	$ 6.50 ☐
	Electronic	$ 13.00 ☐	$ 6.50 ☐
	Print & Electronic	$ 20.00 ☐	$ 10.00 ☐
IEEE Security and Privacy Magazine:	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
	Print	$ 75.00 ☐	$ 37.50 ☐
IEEE Sensors Journal:	Print	$150.00 ☐	$ 75.00 ☐
	Electronic	$ 28.00 ☐	$ 14.00 ☐
Smart Grid, IEEE Transactions on:	Print	$ 50.00 ☐	$ 25.00 ☐
	Electronic	$ 20.00 ☐	$ 10.00 ☐
	Print & Electronic	$ 60.00 ☐	$ 30.00 ☐
Wireless Communications, IEEE Transactions on:	Print	$ 60.00 ☐	$ 30.00 ☐
	Electronic	$ 24.00 ☐	$ 12.00 ☐
	Print & Electronic	$ 60.00 ☐	$ 30.00 ☐
IEEE Wireless Communications Letters:	Print	$ 40.00 ☐	$ 20.00 ☐
	Electronic	$ 9.00 ☐	$ 4.50 ☐
	Print & Electronic	$ 48.00 ☐	$ 24.00 ☐

New! **IEEE Life Sciences Letters** (Open Access Pub) Electronic No Fee
*IEEE membership required or requested
Affiliate application to join SP Society only.   Amount Paid $_____

9.
IEEE Membership Fee (See pricing in Section 8) $_____
Signal Processing Society Fees $_____
Canadian residents pay 5% GST or 13% HST
Reg. No. 125634188 on Society payment(s) & pubs only   Tax $_____
AMOUNT PAID WITH APPLICATION   TOTAL $_____
Prices subject to change without notice.
☐ Check or money order enclosed Payable to IEEE on a U.S. Bank
☐ American Express   ☐ VISA   ☐ MasterCard   ☐ Diners Club

Exp. Date/ Mo./Yr.
Cardholder Zip Code Billing Statement Address/USA Only

Full signature of applicant using credit card   Date

10. WERE YOU REFERRED?
☐ Yes ☐ No   If yes, please provide the follow information:
Member Recruiter Name:_____
IEEE Recruiter's Member Number (Required): _____

IEEE
Signal Processing Society