

IEEE Signal Processing MAGAZINE

Volume 33 | Number 4 | July 2016

MOOC ADVENTURES IN SIGNAL PROCESSING

Acoustic Microphone
Geometry Calibration

Mobile Image Processing
for Health Monitoring

Continuous User Authentication

Application-Oriented Hands-On
Activities for SP Courses



Now... 2 Ways to Access the IEEE Member Digital Library

With **two great options** designed to meet the needs—and budget—of every member, the IEEE Member Digital Library provides full-text access to any IEEE journal article or conference paper in the IEEE *Xplore*® digital library.

Simply choose the subscription that's right for you:

IEEE Member Digital Library

Designed for the power researcher who needs a more robust plan. Access all the IEEE content you need to explore ideas and develop better technology.

- 25 article downloads every month

IEEE Member Digital Library Basic

Created for members who want to stay up-to-date with current research. Access IEEE content and rollover unused downloads for 12 months.

- 3 new article downloads every month

Get the latest technology research.

Try the IEEE Member Digital Library—FREE!

www.ieee.org/go/trymdl



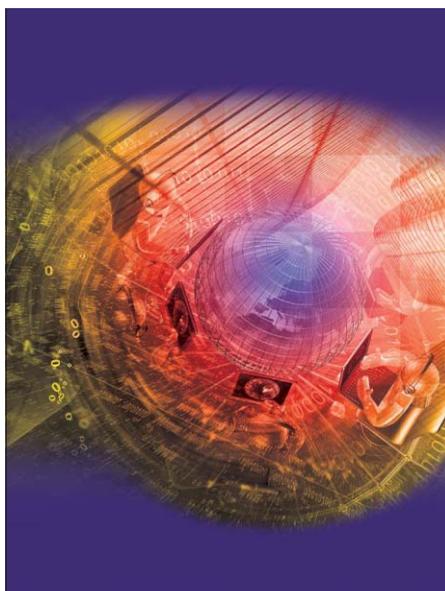
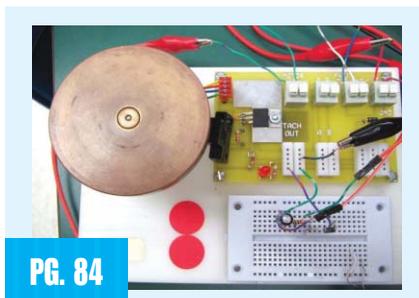
IEEE Member Digital Library is an exclusive subscription available only to active IEEE members.

Contents

Volume 33 | Number 4 | July 2016

FEATURES

- 14 ACOUSTIC MICROPHONE GEOMETRY CALIBRATION**
Axel Plinge, Florian Jacob, Reinhold Haeb-Umbach, and Gernot A. Fink
- 30 SMARTPHONE AND MOBILE IMAGE PROCESSING FOR ASSISTED LIVING**
Hossein Nejati, Victor Pomponiu, Thanh-Toan Do, Yiren Zhou, Sahar Irvani, and Ngai-Man Cheung
- 49 CONTINUOUS USER AUTHENTICATION ON MOBILE DEVICES**
Vishal M. Patel, Rama Chellappa, Deepak Chandra, and Brandon Barbello



ON THE COVER

The last four years have witnessed an explosion of interest and activity in massive open online courses (MOOCs). Online teaching has caused much debate, mostly because of its undeniably disruptive nature with respect to the standard college education paradigm. Check out the article on page 62, as the authors share their collective experiences and insights learned through the design, development, delivery, and management of online DSP courses.

COVER IMAGE: ©ISTOCKPHOTO.COM/ARCHERIX

62 MOOC ADVENTURES IN SIGNAL PROCESSING

Thomas A. Baran,
Richard G. Baraniuk,
Alan V. Oppenheim,
Paolo Prandoni, and
Martin Vetterli

COLUMNS

- 6 Society News**
New Society Officer Elected for 2017 and Nominations Open for 2016 SPS Awards
- 8 Reader's Choice**
Top Downloads in IEEE *Xplore*
- 10 Special Reports**
Signal Processing Plays a Key Role in Wireless Research
John Edwards
- 84 SP Education**
Lessons Learned from Implementing Application-Oriented Hands-On Activities for Continuous-Time Signal Processing Courses
Mario Simoni and Maurice Aburdene
- 90 Life Sciences**
Understanding and Predicting Epilepsy
Christophe Bernard
- 96 Tips & Tricks**
A General Design Method for FIR Compensation Filters in $\Delta\Sigma$ ADCs
Zhe Chen, Shuwen Wang, and Fuliang Yin



IEEE SIGNAL PROCESSING MAGAZINE (ISSN 1053-5888) (ISPREG) is published bimonthly by the Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, 17th Floor, New York, NY 10016-5997 USA (+1 212 419 7900). Responsibility for the contents rests upon the authors and not the IEEE, the Society, or its members. Annual member subscriptions included in Society fee. Nonmember subscriptions available upon request. **Individual copies:** IEEE Members US\$20.00 (first copy only), nonmembers US\$213.00 per copy. Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limits of U.S. Copyright Law for private use of patrons: 1) those post-1977 articles that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923 USA; 2) pre-1978 articles without fee. Instructors are permitted to photocopy isolated articles for noncommercial classroom use without fee. **For all other copying, reprint, or republication permission,** write to IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854 USA. Copyright © 2016 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals postage paid at New York, NY, and at additional mailing offices. **Postmaster:** Send address changes to IEEE Signal Processing Magazine, IEEE, 445 Hoes Lane, Piscataway, NJ 08854 USA. Canadian GST #125634188 **Printed in the U.S.A.**

Digital Object Identifier 10.1109/MSP.2016.2547159

DEPARTMENTS

IEEE Signal Processing Magazine

3 From the Editor

Journey of Learning
Min Wu

4 President's Message

Bringing Signal Processing into the Forefront
of the Conversation
Rabab Ward

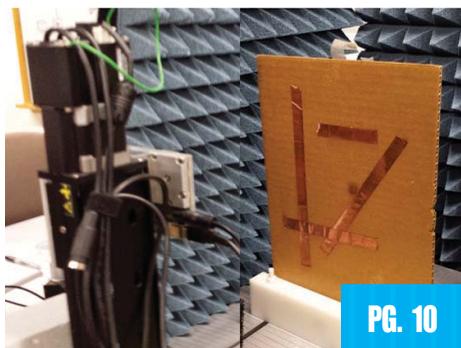
104 Dates Ahead



PG. 104

ICIP 2016 will be held in Phoenix, Arizona, United States,
25-28 September.

©ISTOCKPHOTO.COM/TONDA



PG. 10

EDITOR-IN-CHIEF

Min Wu—University of Maryland, College Park
U.S.A.

AREA EDITORS

Feature Articles

Shuguang Robert Cui—Texas A&M University,
U.S.A.

Special Issues

Douglas O'Shaughnessy—INRS, Canada

Columns and Forum

Kenneth Lam—Hong Kong Polytechnic University,
Hong Kong SAR of China

e-Newsletter

Christian Debes— TU Darmstadt and
AGT International, Germany

Social Media and Outreach

Andres Kwasinski—Rochester Institute
of Technology, U.S.A.

EDITORIAL BOARD

Mritunjoy Chakraborty – Indian Institute of
Technology, Kharagpur, India

George Chrisikos – Qualcomm, Inc.,
U.S.A.

Patrick Flandrin—ENS Lyon, France

Mounir Ghogho—University of Leeds,
U. K.

Lina Karam—Arizona State University, U.S.A.

Hamid Krim—North Carolina State University,
U.S.A.

Sven Lončarić—University of Zagreb, Croatia

Brian Lovell—University of Queensland, Australia

Jian Lu—Qihoo 360, China

Henrique (Rico) Malvar—Microsoft Research,
U.S.A.

Yi Ma—ShanghaiTech University, China

Stephen McLaughlin—Heriot-Watt University,
Scotland

Athina Petropulu—Rutgers University,
U.S.A.

Peter Ramadge—Princeton University,
U.S.A.

Shigeki Sagayama—Meiji University, Japan

Erchin Serpedin—Texas A&M University,
U.S.A.

Shihab Shamma—University of Maryland,
U.S.A.

Hing Cheung So—City University of Hong Kong,
Hong Kong

Isabel Trancoso—INESC-ID/Instituto Superior
Técnico, Portugal

Pramod K. Varshney—Syracuse University,
U.S.A.

Z. Jane Wang—The University of British
Columbia, Canada

Gregory Wornell—Massachusetts Institute
of Technology, U.S.A.

Dapeng Wu—University of Florida, U.S.A.

ASSOCIATE EDITORS—COLUMNS AND FORUM

Ivan Bajic—Simon Fraser University, Canada

Rodrigo Capobianco Guido—
São Paulo State University, Brazil

Ching-Te Chiu—National Tsing Hua University,
Taiwan

Michael Gormish—Ricoh Innovations, Inc.

Xiaodong He—Microsoft Research

Danilo Mandic—Imperial College, U.K.

Aleksandra Mojsilovic—

IBM T.J. Watson Research Center

Fatih Porikli—MERL

Shantanu Rane—PARC, U.S.A.

Saeid Sanei—University of Surrey, U.K.

Roberto Togneri—The University of
Western Australia

Alessandro Vinciarelli—IDIAP-EPFL

Azadeh Vosoughi—University of Central Florida

Stefan Winkler—UIUC/ADSC, Singapore

ASSOCIATE EDITORS—e-NEWSLETTER

Csaba Benedek—Hungarian Academy
of Sciences, Hungary

Paolo Braca—NATO Science and Technology
Organization, Italy

Quan Ding—University of California,
San Francisco, U.S.A.

Pierluigi Failla—Compass Inc, New York,
U.S.A.

Marco Guerriero—General Electric Research,
U.S.A.

Yang Li—Harbin Institute of Technology, China

Yuhong Liu—Penn State University at Altoona,
U.S.A.

Andreas Merentitis—University of Athens, Greece

Michael Muma—TU Darmstadt, Germany

Xiaorong Zhang—San Francisco State University,
U.S.A.

ASSOCIATE EDITOR—SOCIAL MEDIA/OUTREACH

Guijin Wang—Tsinghua University, China

IEEE SIGNAL PROCESSING SOCIETY

Rabab Ward—President

Ali Sayed—President-Elect

Carlo S. Regazzoni—Vice President, Conferences

Konstantinos (Kostas) N. Plataniotis—

Vice President, Membership

Thrasvoulos (Thrasos) N. Pappas—

Vice President, Publications

Walter Kellerman—Vice President,

Technical Directions

IEEE SIGNAL PROCESSING SOCIETY STAFF

Denise Hurley—Senior Manager of Conferences
and Publications

Rebecca Wollman—Publications Administrator

IEEE PERIODICALS MAGAZINES DEPARTMENT

Jessica Barrag , *Managing Editor*

Geraldine Krolin-Taylor, *Senior Managing Editor*

Mark David, *Senior Manager
Advertising and Business Development*

Felicia Spagnoli, *Advertising Production Manager*

Janet Dudar, *Senior Art Director*

Gail A. Schnitzer, Mark Morrissey,

Associate Art Directors

Theresa L. Smith, *Production Coordinator*

Dawn M. Melley, *Editorial Director*

Peter M. Tuohy, *Production Director*

Fran Zappulla, *Staff Director,*

Publishing Operations

Digital Object Identifier 10.1109/MSP.2016.2547160

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.



IEEE prohibits discrimination, harassment, and bullying.
For more information, visit
<http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html>.

FROM THE EDITOR

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

For many of our readers, a career in signal processing may have started from a course, a (text)book, or an article that introduced signal processing to us. With roots in mathematics and physics, signal processing has grown tremendously over the past century; especially after being combined with digital and computing technologies, it has become a vital field that powers our digital life.

One of the missions of *IEEE Signal Processing Magazine (SPM)* is to provide educational value to our community. We appreciate that many readers consider the magazine their first stop in learning about a topic relatively new to them through the magazine's tutorial-style survey and overview articles. This issue of the magazine brings to you several technical features and columns covering a variety of topics.

Advances in multimedia and communications technologies, in which signal processing has played important roles, has enabled new ways to deliver course material beyond traditional textbooks and in-person lectures. Massive open online courses, or MOOCs, became a popular mode of online distance learning a few years ago. *The New York Times* called 2012 "the year of the MOOC," when several well-financed providers emerged with close collaboration with a number of top universities. Courses on these platforms have an impressively high enrollment of learners worldwide. Although the overall participation rate of serious learners who persist to complete course requirements was debatable, one thing is clear: technol-

ogies have changed the ways educational content is presented and shared.

Three groups of signal processing experts from EPFL, the Massachusetts Institute of Technology, and Rice University, widely known for their research and educational contributions to our community, have taken part in creating MOOC educational content on signal processing. In this issue of the magazine, Baran et al. teamed up to reflect on their efforts: What efforts were tried? What worked? What were the commonalities and differences between their MOOCs as well as their on-campus courses? Their special feature article will address these questions.

Going beyond foundational courses, signal and information processing also play an important role in many engineering design projects. Yet there have been few platforms to survey these seemingly adhoc efforts done independently at many institutions. A new article series is being developed by a guest editor team to provide a focused opportunity to share experiences and best practices on undergraduate design projects and hands-on training that incorporates a strong element of signal and information processing. The first article of this series appears in this issue on page 84 and is on application-oriented hands-on activities for learning continuous signal processing. I appreciate the willingness and dedication from the guest editors and the authors to tackle such a topic that is quite different from the usual research articles and is often challenging to put together.

Also, as part of this article series, for the first time, *SPM* is inviting our broad community to share their input on recent

undergraduate design projects related to signal and information processing. The guest editor team will review the submissions to select representative projects based on quality, originality, topic coverage, and diversity, and then they will compile an article to include the highlights of the projects for an upcoming issue of *SPM* with an acknowledgment of project authors and mentors. Please see the announcement on page 102. We look forward to receiving your submissions.

In a broad sense, signal processing education extends beyond degree-oriented education and the walls of universities. Among activities supported by the IEEE Signal Processing Society (SPS) are popular tutorial programs in major conferences; Chapter events and seasonal schools that run multiple times a year worldwide; a signature series of invited talks on technical trends and overviews at SPS's outreach effort of ChinaSIP and the Signal and Data Science Forum, and more. Now, teaching notes and lecture slides are also hosted on the IEEE SigPort repository, and videos and webinars of tutorials and technical talks can be found on the SigView resource center.

According to an old saying, it is never too late to learn. Whether you prefer the old-style learning or are open to trying some fashionable new ways, I hope you will find *SPM* and other mechanisms provided by SPS a trusted source to help your lifelong journey of learning signal processing.

SP

PRESIDENT'S MESSAGE

Rabab Ward | SPS President | rababw@ece.ubc.ca

Bringing Signal Processing into the Forefront of the Conversation

The general public's lack of understanding about signal processing and its applications has been a continued hurdle for our Society, our members, and those in our field. Despite its undeniable pervasiveness throughout our daily lives, signal processing is largely invisible. Its effects, however, are apparent in our everyday interactions with technology—in mobile devices, in the booming wearable industry, in forensics, in national security, and so much more. While our field enables and empowers hundreds of tangible devices and applications, the signal processing behind those devices and applications remains largely unknown. How do we, as a community, generate awareness about an invisible field that is integral in enabling highly visible commercial products and applications? How do we make the invisible visible?

This is an issue that the IEEE Signal Processing Society (SPS) has been trying to tackle for years, and over the past year and a half, we've been making a concentrated effort to bring signal processing into the forefront of the conversation. This year will be an exciting one for our Society—we will be launching a brand new website that will be a centralized resource for the SPS and signal processing disciplines as a whole. In addition to receiving a refreshing facelift and revitalized content for members and volunteers, the new

SPS website will feature content intended to teach nontechnical demographics and answer the question, in layman's terms, "What is signal processing?"

The new website will be a jumping-off point for a multifaceted outreach effort to make signal processing approachable to a larger, nontechnical audience and inspire future generations of engineers to consider signal processing as a viable area of study and career path. With the assistance of a public relations firm, we will conduct media outreach efforts to promote signal processing and its applications across commercial publications, trade publications, and blogs. In doing so, we will establish SPS volunteers and members as leaders in the field, providing opportunities for them to contribute their expertise to external publications to reach previously untapped audiences.

We will be building upon our internal content too, ensuring that we don't lose sight of what's important to our diverse and evolving member base. In addition to *IEEE Signal Processing Magazine* and the *Inside Signal Processing Newsletter*, we will be launching a new SPS blog that will be home to SPS news and educational content contributed by our volunteers. If you're interested in contributing content to the SPS blog, or have other questions about the SPS visibility initiative, contact SPS Membership and Content Administrator Jessica Perry at jessica.perry@ieee.org.

More introductory videos are being created by our technical committees. Our first two videos, "What Is Signal Processing?" and "Signal Processing and Machine Learning," were received very favorably and were recently translated into Mandarin, Arabic, and Spanish. All of these videos can be accessed on our YouTube channel; visit https://www.youtube.com/channel/UCAgRLbgh-cW_8U_JN97RsyA. We plan to continue making videos that present signal processing concepts in an interesting and engaging way and promoting them to members and the general public alike.

The task of promoting signal processing and its applications also falls into the hands of our members—academics, industry professionals, and researchers—who are working and living in it every single day. For that reason, we will be providing a host of membership development tools, Chapter tool kits, brochures, and additional resources to educate and engage our target audiences about signal processing and the benefits of SPS membership.

It is a pivotal and exciting time to be an SPS member, and we look forward to sharing all of these items with you and laying the foundation for the next era of the SPS.

SP

Digital Object Identifier 10.1109/MSP.2016.2566718
Date of publication: 1 July 2016



IEEE International Conference on Acoustics, Speech, and Signal Processing

March 5-9, 2017 New Orleans, USA

Internet of Signals

Preliminary Call for Papers

As music and rhythm are the heartbeats of life, signal and information processing is the heartbeat of IT technology development for our daily life. Having both of them capture the hearts and souls of all the attendees of the 42th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017) will be held in Hilton Conference Centre, at the Jazz music capital (New Orleans, USA) on March 5-9, 2017. ICASSP is the world's largest and most comprehensive technical conference focused on signal processing and its applications. The conference provides, both for researchers and developers, an engaging forum to exchange ideas and proposed new developments in this field. The theme of the ICASSP 2017 is "The Internet of Signals" which is the real technology and world behind the Internet of Things. The conference will feature world-class International speakers, tutorials, exhibits, lectures and poster sessions from around the world. Topics include but are not limited to:

- Audio and acoustic signal processing
- Bio-imaging and biomedical signal processing
- Design & implementation of sig. processing sys.
- Image, video & multidimensional sig processing
- Industry technology tracks
- Information forensics and security
- Machine learning for signal processing
- Multimedia signal processing
- Remote Sensing and signal processing
- Signal Processing for Smart Systems
- Sensor array & multichannel signal processing
- Signal processing education
- Signal processing for comm. & networking
- Signal processing theory & methods
- Signal processing for Big Data
- Internet of Things and RFID
- Speech processing
- Spoken language processing
- Signal Processing for Brain Machine Interface
- Signal Processing for Cyber Security

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 papers will be made by the ICASSP 2017 committee upon recommendations from Technical Committees.

Special Session Proposals:

Special session proposals should be submitted by July 11th, 2016. Proposals for special sessions must include a topical title, rationale, session outline, contact information for the session chair, a list of authors, and a tentative title and abstract. Additional information can be found at the ICASSP 2017 website (www.ieee-icassp2017.org).

Tutorials:

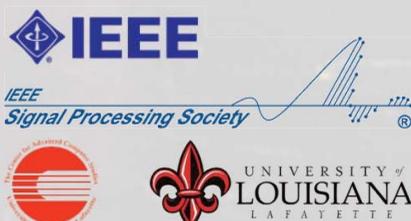
Will be held on March 5th, 2017. Brief proposals should be submitted by July 11th, 2016. Proposal for tutorials must include a title, an outline of the tutorial and its motivation, a two-page CV of the presenter(s), and a short description of the material to be covered.

Signal Processing Letters:

Authors of IEEE Signal Processing Letters (SPL) papers will be given the opportunity to present their work at ICASSP 2017, subject to space availability and approval by the ICASSP Technical Program Chairs. SPL papers published between January 1, 2016 and December 31, 2016 are eligible for presentation at ICASSP 2017. Because they are already peer-reviewed and published, SPL papers presented at ICASSP 2017 will neither be reviewed nor included in the ICASSP proceedings. Requests for presentation of SPL papers should be made through the ICASSP 2017 website on or before December 12th, 2016. Approved requests for presentation must have one author/presenter register for the conference.

Demos:

Offers a perfect stage to showcase innovative ideas in all technical areas of interest at ICASSP. All demo sessions are highly interactive and visible. Please refer to the ICASSP 2017 website for additional information regarding demo submission.



Important Deadlines:

Tutorials & Special session proposals:	July 11 th , 2016
Notification of special session & tutorial acceptance:	August 15 th 2016
Submission of regular papers:	September 12 th , 2016
Signal Processing Letters:	November 21 st , 2016
Notification of paper acceptance:	December 12 th , 2016
Author registration:	January 9 th , 2017

SOCIETY NEWS

New Society Officer Elected for 2017 and Nominations Open for 2016 SPS Awards

The Board of Governors (BoG) of the IEEE Signal Processing Society (SPS) elected one new officer who will start his term on 1 January 2017: Nikos Sidiropoulos (University of Minnesota) will serve as 2017–2019 SPS vice president-Membership. He succeeds Kostas Plataniotis (University of Toronto), who has held the post of vice president-Membership since January 2014.

Nikos Sidiropoulos: New Society Officer



Nikos Sidiropoulos received the diploma in electrical engineering from the Aristotelian University of Thessaloniki, Greece, and the M.S. and Ph.D.

degrees in electrical engineering from the University of Maryland–College Park, in 1988, 1990, and 1992, respectively. He has served as assistant professor at the University of Virginia (1997–1999); associate professor at the University of Minnesota (2000–2002); professor at the Technical University of Crete, Greece (2002–2011); department chair (2005–2007); and professor at the University of Minnesota since 2011, where he holds the ADC Chair in Digital Technology. His research interests are in signal

processing, wireless communications, optimization, and tensor decomposition. His current research focuses on signal and tensor analytics for learning from big data. He received the National Science Foundation/CAREER Award in 1998 and the IEEE SPS Best Paper Award in 2001, 2007, and 2011. He was an IEEE SPS Distinguished Lecturer (2008–2009); chair, SPS Signal Processing for Communications and Networking Technical Committee (2007–2008); and has held several associate editor positions for SPS journals. He received the 2010 IEEE SPS Meritorious Service Award and the 2013 Distinguished ECE Alumni Award from the University of Maryland. He is a Fellow of the IEEE (2009) and of EURASIP (2014).

Nominations open for 2016 SPS awards

The SPS Awards Board is now accepting nominations for all Society-level awards, from paper awards to the major Society awards. Nominations are due by 1 September 2016 and should be submitted to Theresa Argiropoulos (t.argiropoulos@ieee.org), who will collect the nominations on behalf of Awards Board Chair José M.F. Moura. Nominators should take into consideration the need for representation of diversity in the nomination slate when submitting their nominations. Full details on the nomination process, details on each award, as well as the nomination and endorsement forms, are

available on the Society's website: <http://www.signalprocessingsociety.org/awards-fellows/awardspage/>.

- *Who can nominate:* Nominations are accepted from any Society individual member, Society committee, or Society board. Nominations from individual members can be supported by up to two endorsement letters from two other individual members.
- *Which awards:* Each year, the SPS honors outstanding individuals who have made significant contributions related to signal processing through the Society Award, the Industrial Leader Award, the Industrial Innovation Award, the Technical Achievement Award, the Education Award, and the Meritorious Service Award. The Society also recognizes outstanding publications in SPS journals and magazines through the Best Paper Award, Overview Paper Award, Sustained Impact Paper Award, IEEE Signal Processing Letters Best Paper Award, IEEE Signal Processing Magazine Best Column Award, IEEE Signal Processing Magazine Best Paper Award, and the Young Author Best Paper Award.

Nominations for the Best Paper Award and Young Author Best Paper Award should refer to the papers published in the following Society journals:

- *IEEE Journal of Selected Topics in Signal Processing (JSTSP)*
- *IEEE Transactions on Audio, Speech, and Language Processing (T-ASLP)*

- *IEEE Transactions on Image Processing (T-IP)*
- *IEEE Transactions on Information Forensics and Security (T-IFS)*
- *IEEE Transactions on Signal Processing (T-SP)*.

SPS awards procedural changes

Over the past few years, the Society has approved some procedural changes to the SPS Awards program, including some new changes approved earlier this year. Please note that these changes are in effect for the 2016 nomination period. The changes are intended to provide an effective means to encourage award nominations in all categories from the SPS community-at-large, including individuals, technical committees, editorial boards, and other major boards, except in the cases of conflict of interests. Technical committees and boards may pass on to the Awards Board one or multiple nominations that they receive for all awards.

The paper awards nomination form now requests citation impact information.

The Awards Board will continue to review the nominations and make selections on paper awards.

For all major awards other than paper awards, the Awards Board will now be responsible for vetting the nominations and producing a short list of no more than three nominations per award. The Board of Governors will continue to vote on the selection of the major awards.

A board or committee cannot nominate one of its current members for an award. However, the board/committee member can be nominated by another board or committee. Current elected members of a committee/board may participate as individual nominators for other members of the same board/committee. Individual nominations can have multiple conominators listed on the nomination form. In addition, individual nominations can include up to two endorsements to strengthen the nomination. All endorsements must be submitted via e-mail to the specified address, which will provide the nomination with a date and time stamp. If more than two

endorsements are submitted, only the first two received endorsements will be forwarded to the IEEE SPS Awards Board for consideration. A nominator cannot serve as an endorser for a nomination he/she is submitting. If the Society policies state that a particular board/committee/individual is not eligible to nominate for a particular award, then that same group of individuals is not eligible to be an endorser.

Technical committee and special interest group award nomination procedures have been approved with suggested award nomination and voting procedures.

For full details on each award, as well as the new Society and technical committee/special interest group awards policies and endorsement form, please visit the Society's website at <http://www.signal-processing.org/awards-fellows/awardspage/>.

If you have any questions regarding the process, please contact Awards Board Chair José M.F. Moura at moura@ece.cmu.edu.



SigPort.org

Do you know? Your colleagues are archiving slides of their signal processing work on IEEE SigPort.

The slides and posters you spent hours to make are highlights of your work. Aren't they "forgotten" soon after conference presentations or thesis defense?

IEEE Signal Processing Society's SigPort repository helps extend the life of your slides and posters, and raise the visibility of your work. SPS Members upload FREE in 2016!

- **Promote your work more and sooner than IEEE Xplore:** ICASSP'16 slides and posters posted on SigPort received an average of 32 downloads within two months. GlobalSIP'15 slides had an average of 91 downloads within six months.
- **How?** Login on www.sigport.org using IEEE web account credentials. Go to "submit your work" on the top menu and use promotion code **you14100** for free upload.
- **Beyond slides and posters:** SigPort welcomes research drafts, white papers, theses, slides, posters, lecture notes, dataset descriptions, product brief, and more. Send questions or comments to Chief Editor at [<yansun@uri.edu>](mailto:yansun@uri.edu).



Digital Object Identifier 10.1109/MSP.2016.2583385

problem common in popular graphical password systems.

June 2014

Catch Me If You Can: Evaluating Android Anti-Malware Against Transformation Attacks

Rastogi, V.; Chen, Y.; Jiang, X.

The authors developed DroidChameleon, which transforms Android apps in various ways. Tests with ten popular commercial anti-malware applications show that none are resistant to all transformations, and most are defeated by trivial transformations. The authors propose remedies for improved malware detection.

January 2014

Contrast Enhancement-Based Forensics in Digital Images

Cao, G.; Zhao, Y.; Ni, R.; Li, X.

Creating believable image forgeries requires a consistent contrast across the image, such as when image parts are pasted, local contrast enhancement is often applied. This paper detects global contrast enhancement and variances between regions. The authors use statistics of the JPEG-compressed images for the detection.

March 2014

Thermal Facial Analysis for Deception Detection

Rajoub, B.A.; Zwigelaar, R.

The authors use thermal imaging to monitor the periorbital region's thermal variations and test whether it can offer a discriminative signature for detecting deception. They report an 87% ability to predict the lie/truth responses based on a within-person methodology, but between person modeling did not generalize across the test set.

June 2014

Designing an Efficient Image Encryption-Then-Compression System via Prediction Error Clustering and Random Permutation

Zhou, J.; Liu, X.; Au, O.C.; Tang, Y.Y.

The proposed image encryption scheme operated in the prediction error domain



©ISTOCKPHOTO.COM/KALAWIN

is shown to be able to provide a reasonably high level of security. We also demonstrate that an arithmetic coding-based approach can be exploited to efficiently compress the encrypted images.

January 2014

Reversible Data Hiding in Encrypted Images by Reserving Room Before Encryption

Ma, K.; Zhang, W.; Zhao, X.;

Yu, N.; Li, F.

All previous methods embed data by reversibly vacating room from the encrypted images, which may be subject to some errors on data extraction and/or image restoration. In this paper, we propose a novel method by reserving room before encryption with a traditional reversible data hiding (RDH) algorithm, and thus it is easy for the data hider to reversibly embed data in the encrypted image.

March 2013

Data Hiding in Encrypted H.264/AVC Video Streams by Codeword Substitution

Xu, D.; Wang, R.; Shi, Y.Q.

In this paper, a novel scheme of data hiding directly in the encrypted version of H.264/AVC video stream is proposed, which includes three parts: H.264/AVC video encryption, data

embedding, and data extraction. By analyzing the property of H.264/AVC codec, the codewords of intraprediction modes, the codewords of motion vector differences, and the codewords of residual coefficients are encrypted with stream ciphers. Then, a data hider may embed additional data in the encrypted domain by using codeword substitution technique.

April 2014

Permission Use Analysis for Vetting Undesirable Behaviors in Android Apps

Zhang, Y.; Yang, M.; Yang, Z.;

Gu, G.; Ning, P.; Zang, B.

This paper presents VetDroid, a dynamic analysis platform for generally analyzing sensitive behaviors in Android apps from a novel permission use perspective. VetDroid proposes a systematic permission use analysis technique to effectively construct permission use behaviors, i.e., how applications use permissions to access (sensitive) system resources, and how these acquired permission-sensitive resources are further utilized by the application.

November 2014

SP

Signal Processing Plays a Key Role in Wireless Research

New technologies promise to boost speed, bandwidth, availability, and reliability

As the world grows increasingly dependent on wireless connections, existing untethered communication technologies are being stretched to the limit. The growing demands of businesses, governments, and consumers for ever greater wireless bandwidth pose a major challenge for a wide range of researchers, who face the unenviable task of developing enhanced wireless technologies and approaches as quickly as possible. In many areas, signal processing is playing an important role in helping researchers meet their goal.

Riding the millimeter wave

Wireless devices, including everything from smartphones to headphones, are now an integral part of everyday life. Yet as the number of wireless devices multiplies, technology developers and users face increasing bandwidth and spectrum limitations, impacting device usability, availability, and reliability. The rapidly emerging Internet of Things (IoT) will inevitably place even more pressure on available wireless resources as untold numbers of sensors arrive to interconnect virtually everything—and possibly everybody—on the planet.

Researchers worldwide are fighting back against an accelerating avalanche of wireless bottlenecks by developing innovative technologies that aim to boost data transfer speeds, increase range, and

conserve limited spectrum space. “Mobile broadband data demand will grow by a 1,000-fold between 2010 and 2020,” predicts Upamanyu Madhow, a professor of electrical and computer engineering at the University of California, Santa Barbara (UCSB). “A natural approach to meeting this demand is to employ small cells with millimeter-wave links from base stations to mobiles, as well as millimeter-wave links for wireless backhaul from base stations to the optical backbone,” he says.

According to Madhow, millimeter-wave communication represents the future of wireless communication,

offering the potential to provide data rates approaching optical speeds. Possible applications include indoor wireless networks and next-generation cellular networks. Low-cost, short-range millimeter-wave radar systems may also eventually play a significant role in the development of applications such as situational awareness for autonomous and semiautonomous vehicles as well as gesture recognition for handheld devices (Figure 1).

Madhow and fellow UCSB researchers Jim Buckwalter, Mark Rodwell, and Heather Zheng, along with Amin Arbabian of Stanford University and Xinyu

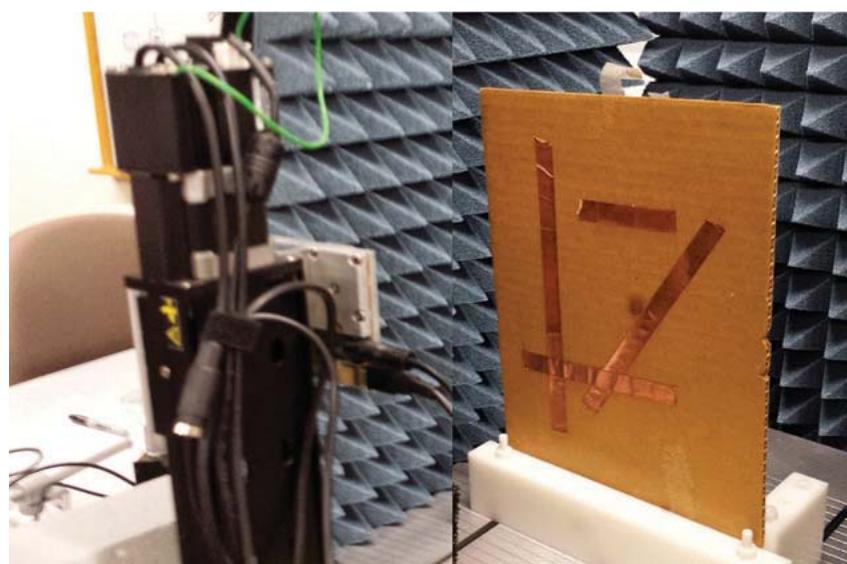


FIGURE 1. A test bed for improved gesture recognition is one part of a multifaceted millimeter-wave research project being conducted at UCSB, Stanford University, and the University of Wisconsin–Madison.

Digital Object Identifier 10.1109/MSP.2016.2547161
Date of publication: 1 July 2016

Zhang of the University of Wisconsin–Madison, plan to demonstrate the feasibility of a large-scale millimeter-wave wireless data network that can operate at gigabit speeds. Their goal is to develop a technology using concepts that span signal processing, network protocols, circuit design, and communication architecture. “The highly directional nature of millimeter-wave links requires rethinking wireless network protocols based on omnidirectional transmission, such as the ‘listen before talk’ protocol underlying Wi-Fi,” Madhow says.

“Signal processing is at the heart of both the communications and radar research we are pursuing,” Madhow says. The hardware characteristics necessary for millimeter-wave systems are quite different from those in existing wireless systems, which calls for the invention of new signal processing techniques. “For example, the small wavelengths make it possible to realize antenna arrays with a very large number of elements in a small form factor—a 1,000-element antenna array could fit within the palm of the hand—which makes it possible to synthesize narrow, electronically steerable beams.”

Such arrays typically use radio-frequency (RF) beamforming that allows control over the phases—and, in some cases, the amplitudes—of each array element but does not allow access to the signals transmitted or received by the individual elements. “This means that standard adaptive signal processing techniques, typically based on least squares, cannot be applied to train such arrays,” Madhow says. “We have therefore had to invent new compressive techniques for this purpose.”

The biggest signal processing challenge facing millimeter-wave communication and radar system developers is the fact that classical algorithms and models often do not apply to the technology, due to its tiny wavelengths and associated hardware constraints. “Thus, while we can leverage many core concepts, we have had to invent new approaches both in terms of modeling and algorithms,” Madhow says.

To allow millimeter-wave signals to work with relatively large and complex

antenna arrays, the researchers developed a compressive estimation approach. The technique exploits the sparsity of the millimeter-wave channel to estimate, via randomized measurements, the directions of the dominant arrival/departure paths between the transmitter and receiver. The approach falls into the general area of sparse modeling and compressive sensing, Madhow says. “However, standard compressive sensing algorithms do not apply to the settings of interest to us, since they are based on sparsity in a discrete setting, while the parameters we wish to estimate are continuous-valued.”

Although a substantial amount of theory already exists in the area of compressive sensing, the researchers still needed to develop a general theory of compressive estimation. “Over the past few years, we have developed both theories and algorithms in this area,” Madhow says. “Our goal now is to experimentally demonstrate these using a millimeter-wave test bed and to design network architectures and protocols around these core signal processing ideas.”

Other areas of millimeter-wave research also required fresh approaches. For short-range millimeter-wave imaging—used in applications such as gesture recognition and vehicular situational awareness—the researchers realized that signal processing must be based on models that go beyond the classic point scatterer target model, since targets appear larger at short ranges and small wavelengths. “We are therefore pursuing new patch-based target models optimized using tools from estimation theory,” Madhow says. “We are employing both classical correlation-based signal processing and sparse signal processing techniques based on these models.”

For handling large communication bandwidths, the researchers have explored several different options, trading off both analog and digital signal processing. Two recent examples are a

“mostly analog” processing architecture for a line-of-sight (LoS), multiple-input, multiple-output (MIMO) prototype, and an analog multiband approach for dividing the communication bandwidth into smaller slices that can be efficiently digitized for processing.

“Modern communication transceiver design is heavily based on digital signal processing (DSP) and can therefore exploit the economies of scale from Moore’s law,” Madhow observes. “This

approach is predicated on analog-to-digital converters (ADCs) that can faithfully represent signals in the digital domain,” he says. “As we increase the communication bandwidth, however, ADCs

become costly and power-hungry, hence we have been exploring various alternatives that trade off the complexity of analog and DSP.”

Biomimetic antenna arrays

Engineers at the University of Akron’s College of Engineering’s Wireless Communications Lab are developing electrically small biomimetic antenna arrays inspired by one of the most sensitive auditory systems in the natural world: an insect’s ear system. In collaboration with the University of Wisconsin–Madison, researchers are investigating methods aimed at increasing the data rate in MIMO wireless communication systems while also reducing the size and power consumption of associated mobile devices.

“The overall objective of this interdisciplinary research project is to use recent advances in the areas of multiantenna wireless communications, signal processing, and coupled antenna array (CAA) technology to enhance the efficiency of spectrum utilization of mobile wireless communication systems,” says Hamid Bahrami, the project’s principal investigator and an associate professor in the University of Akron’s Department of Electrical and Computer Engineering. The project’s coprincipal investigator is Nader Behdad at the University of

“Signal processing is at the heart of both the communications and radar research we are pursuing.”

Wisconsin–Madison. “My focus will be on the signal processing and communications aspect, and Dr. Behdad’s focus will be on the antenna science aspect of the project,” Bahrami says.

The project’s inspiration is based on the fact that many small animals and insects possess a hyper-acute sense of directional hearing. “In particular, the parasitoid fly *Ormia ochracea* (Figure 2) demonstrates one of the most sensitive auditory systems in the animal world and is capable of detecting the direction of arrival (DoA) of a sound wave with a 1–2° angular resolution,” Bahrami says. The fly’s unique “ears” are complex structures located inside the fly’s chest, near the base of its front legs. The fly is too small for the time difference of sound arriving at the two ears to be calculated in the usual way, yet it can determine the direction of sound sources with high precision.

Current-generation mobile devices, such as 4G smartphones, typically incorporate two or four antennas to

increase performance. Although a multiple antenna design will usually boost communication speed and lower a device’s power consumption, the approach also requires extra space, which inevitably leads to larger device form factors. Since the distance between antenna array elements can cover multiple inches, depending on the frequency

band, a smartphone radio containing more than four antennas would be prohibitively large. Using *Ormia ochracea* as their model, the researchers aim to shrink the spacing between MIMO antennas inside smart-

phones, allowing extra antennas to be added without sacrificing portability.

“In the short term, the expected research findings will lead to the development of multiantenna receivers with offered spectral efficiencies that cannot be easily achieved using any other existing technology today,” Bahrami says. “In the long term, these concepts are expected to revolutionize how we solve problems in a variety of areas ranging from

high-capacity MIMO systems and multiantenna RF front-end architectures to small-aperture, high-resolution microwave/millimeter-wave radar and imaging systems.”

The researchers are using statistical signal processing as well as detection theory to optimally detect the received signals over CAA elements. “We also use information theory to analyze the achievable capacity of CAAs,” Bahrami says.

“To fully benefit from the nature of coupling in CAAs, it is important to apply advanced signal processing techniques for detection,” Bahrami notes, adding that conventional approaches might not work since they are designed for traditional antenna arrays, such as phased array antennas.

To achieve the best signal processing techniques for signal detection, the researchers need to first develop detailed signal-level and system-level models for the CAAs. This looms as a challenging task, due to the CAAs relatively complex architecture. “Then, when it comes to capacity analysis, we should use the developed system-level model,” Bahrami says. “Due to the limited number of information theoretic tools available, this is also one of the main challenges that needs to be addressed.”

The researchers’ main challenge beyond signal processing is altering the CAAs’ characteristics to improve the capacity of MIMO communications systems. “The original CAA design targets high directivity,” Bahrami says. “We need to modify that in a way that provides a higher capacity.” Coprincipal investigator Behdad is working to develop suitable coupling networks that target increased capacity, Bahrami says.

Another challenge is increasing the CAAs’ operational bandwidth. The researchers’ preliminary CAA designs, utilizing coupled monopole antennas, have typical bandwidths of 50–100 MHz at 1 GHz. “While we anticipate such bandwidths to be sufficient for many wireless systems, in some applications larger bandwidths may be required,” Bahrami says.

The main challenge beyond signal processing is altering the CAAs’ characteristics to improve the capacity of MIMO communications systems.



FIGURE 2. The ear system of a parasitoid fly has acute directional sensitivity, an attribute that is inspiring antenna research at the University of Akron and the University of Wisconsin–Madison.

“We have shown that our improved CAA design can increase the capacity compared to conventional arrays by a factor of 20–30% at small antenna spacings,” Bahrami says. “The biggest challenge now is to go even beyond this.”

Underwater wireless

Approximately two-thirds of the Earth’s surface is covered by water, and it is inevitable that the number of underwater sensors will grow significantly over the next few years. Yet data retrieval from submerged sensors continues to be hindered by the limited speed of existing underwater communication networks, which remain steadfastly stuck somewhere in the range of late 20th-century dial-up modems. This stubborn obstacle hampers a wide range of activities, including search-and-rescue operations, pollution monitoring, tsunami detection, and other types of important work conducted in bodies of waters such as oceans, bays, rivers, and lakes.

State University of New York at Buffalo (SUNY-Buffalo) researchers are now working to help underwater communications at least partially catch up with over-the-air data transmission rates. “The remarkable innovation and growth we have witnessed in land-based wireless communications has not yet occurred in underwater sensing networks, but we are starting to change that,” says Dimitris Pados, a professor of electrical engineering in the SUNY-Buffalo’s School of Engineering and Applied Sciences.

Pados and several coresearchers are developing new types of hardware and software, including waterproof modems and open-architecture protocols, designed to address underwater transmission speed issues. The teams’ efforts are currently focused on combining a software-defined radio (SDR) with underwater acoustic modems. Their prototype underwater communication system is a cognitively self-optimized technology that works by jointly adapting link signal waveforms and network session routes to maximize



FIGURE 3. SUNY-Buffalo students test an enhanced underwater communications technology on Lake Erie.

network throughput and/or spectral efficiency under a wide range of operating conditions.

“Link waveforms are cognitively optimized over the whole spectrum accessible by the transceiver nodes (all-spectrum channelization) and routes are cognitively optimized over all accessible nodes,” Pados says. He notes that the technology is based on well-understood signal processing theory. “An elementary pulse signal is selected and shaped to occupy all hardware accessible frequency bandwidth—raised cosine pulse, chirps, and others,” Pados explains. A finite number of different sign/phase shifted repeats of this pulse forms the final waveform design, which will then carry the information symbols on its back. “The specific values of the sign/phase shift sequence—that we will call waveform-carrier code—are determined near-real time by principal-component signal analysis techniques that calculate the code with highest interference avoidance properties at any given time and receive node location,” Pados says.

State University of New York at Buffalo researchers are now working to help underwater communications at least partially catch up with over-the-air data transmission rates.

In May 2015 at Lake LaSalle on the SUNY-Buffalo campus, the research team demonstrated its technology for the first time, achieving communication at 200,000 bits/s over an underwater distance of 200 m. The system consisted of two Ettus Research USRP N210 SDN modules using Teledyne RESON TC4013 transducers with a 1 Hz–170 kHz operational frequency range. Further tests were conducted on nearby Lake Erie (Figure 3).

Pados says he expects to see many significant applications for the technology emerge over the next few years as it matures and achieves increasingly faster throughput speeds. Real-time underwater sensing (pollution, temperature, or sea currents), sea-life monitoring, port surveillance, wireless diver-to-diver communication, wireless diver/underwater vehicle communication, untethered sea exploration, search-and-rescue operations, underwater wireless video feeds, and off-shore drilling monitoring are just some of the technology’s potential applications, he notes. **SP**



Axel Plinge, Florian Jacob,
Reinhold Haeb-Umbach,
and Gernot A. Fink

Acoustic Microphone Geometry Calibration

An overview and experimental evaluation of state-of-the-art algorithms

Today, we are often surrounded by devices with one or more microphones, such as smartphones, laptops, and wireless microphones. If they are part of an acoustic sensor network, their distribution in the environment can be beneficially exploited for various speech processing tasks. However, applications like speaker localization, speaker tracking, and speech enhancement by beamforming avail themselves of the geometrical configuration of the sen-

sors. Therefore, acoustic microphone geometry calibration has recently become a very active field of research. This article provides an application-oriented, comprehensive survey of existing methods for microphone position self-calibration, which will be categorized by the measurements they use and the scenarios they can calibrate. Selected methods will be evaluated comparatively with real-world recordings.

Introduction

Wireless acoustic sensor networks (WASNs) are a promising approach for sound capturing and processing systems [2].

Digital Object Identifier 10.1109/MSP.2016.2555198
Date of publication: 1 July 2016

With the low cost of acoustic sensors and wireless communication devices, such networks are more common. They are composed of devices such as smartphones, tablet computers, wireless microphones, or hearing aids that are equipped with a single or multiple microphones. Due to their distribution in an environment, it is likely that at least one device is close to every relevant sound source. Thus, WASNs deliver a signal with improved quality compared to traditional microphone arrays, which sample a sound field only locally. As a consequence of the ad hoc nature of many WASNs, the position of the sensor nodes is often unknown and may even vary over time. However, a number of important speech processing tasks rely on the estimation of the location of sound sources, which in turn requires the location of the recording devices to be known.

A popular application of distributed microphones or microphone arrays is source localization and tracking. The localization results are used to enable subsequent applications, such as camera control and speech enhancement. But only if the positions and orientations of the microphones are known can the source position or direction be estimated by triangulation, trilateration, or other approaches. Errors in the assumed geometric arrangement of the microphones have a significant effect on the localization accuracy, as is well demonstrated by the experiment described in “Impact of Geometry Errors on Source Localization.”

While signal extraction from distributed microphone arrays can be achieved without explicit estimation of the position of the sources [22], the speech enhancement performance can be improved by incorporating source location information. Faster adaptation to changing acoustic environments can be obtained by so-called informed spatial filtering approaches, where the adaptation of the source extraction filters is supported by information on the source location [42]. Moreover, in parametric spatial processing, parameters describing the sound field, such as the source location, are employed for spatial audio coding and reproduction, source enhancement, or acoustic scene analysis.

Given the importance of the aforementioned speech processing tasks, which require the geometric configuration of the acoustic sensors to be known, research in microphone geometry calibration has substantially increased in recent years. While, in early approaches, the positions of the microphones were determined by hand or computed from manual measurements of pairwise microphone distances, this clearly becomes impractical in the ad hoc scenarios typical of WASNs.

The goal of recent research efforts is to devise methods to infer the position of the sensors solely from the acoustic signals they capture, a strategy termed *acoustic geometry calibration*. Research articles also refer to the task as *microphone self- or auto-localization*, or *position self-calibration*. The basic idea is to extract a quantity from the microphone signals that is related to their geometric arrangement—for example, the time difference of an arriving sound at two nodes, or the direction under which an acoustic event is observed. The extracted information

is used in an objective function that essentially scores the deviation of the actual measurement from the measurement as predicted by the assumed geometry. Since the extraction of location information from the microphone signals is often achieved by measurement of time or time difference, there exists a close dependence of localization on clock synchronization, as we will see in the following.

We also consider node localization in nonacoustic sensor networks. However, there are some fundamental differences. Popular localization systems, such as satellite, cellular, or Wi-Fi-based systems, rely on knowledge of the location of anchor nodes, i.e., the satellites or base stations that transmit a radio signal, to infer the position of user terminals. Such anchor nodes are, in general, not available in acoustic geometry calibration. Furthermore, the nodes in wireless sensor networks are typically assumed to consist of a transceiver, i.e., a radio transmitter and a receiver at the same location. Thus, active localization can be performed by exchanging time stamps or other signaling information, from which position-related information is estimated.

In an acoustic sensor network, we wish to perform localization via acoustic signals only. There are also active approaches, which can be used for sensor nodes such as smartphones or laptop computers equipped with both loudspeakers and microphones. In the general case of passive calibration, however, the sensor is unable to produce a sound by itself, and the process has to rely on external acoustic events such as speech or ambient noise. Then there is no time synchronization between transmitter and receiver, and time or time difference estimation is further complicated by the unfavorable correlation properties of the signals.

Only recently has this unconstrained scenario been tackled by acoustic geometry calibration algorithms. Earlier approaches targeted laboratory installations and required a known arrangement of loudspeakers, dedicated calibration signals, and strict time synchronization. By removing the earlier constraints, modern approaches attempt to calibrate ad hoc arrangements using ambient sounds such as speech. Furthermore, the constraints of time synchronization or collocation of microphones and loudspeakers are relaxed.

This ongoing research is being performed in a number of directions, focusing on different scenarios and employing different measurements and optimization strategies. The purpose of this article is to categorize the individual approaches with respect to the scenario addressed and the signals and objective functions employed. This should give the reader a basic understanding of the algorithms used, their applicability to different scenarios, and the expected localization performance.

Application scenarios

Acoustic geometry calibration approaches have been developed for different application scenarios. In the following survey, we distinguish three fundamental types of microphone arrangements: small compact arrays, distributed individual microphones, and distributed microphone arrays

Impact of Geometry Errors on Source Localization

To investigate the dependence of acoustic source localization accuracy on the microphone position calibration performance, a recent speaker tracking method [32] was applied to simulated data. Five nodes with circular microphone arrays composed of five microphones each were used, located in the middle of the room. A speaker was localized at 18 positions around the arrays in a reverberant room ($T_{60} = 0.5$ seconds) of size $6.5 \text{ m} \times 3.5 \text{ m} \times 2.5 \text{ m}$. An erroneous geometry calibration was simulated as follows: the calibration error was drawn from a zero mean normal distribution with increasing standard deviation. This way, a fixed mean calibration error from 0.1 m and 2° to 0.5 m and 10° over all arrays was simulated. The plots in

Figure S1 show the mean speaker localization error $\varepsilon_l(m)$ as a function of mean position [Figure S1(a)] and orientation [Figure S1(b)] geometry calibration error. Furthermore, the error bars indicate the standard deviation over 100 experiments. The localization error should be no larger than the size of a human head, i.e., below 30 cm (indicated by the orange line), for practical applications such as camera control. Due to reverberation deteriorating the measurements, the resulting localization error is already close to that level even without any calibration error. When the calibration error exceeds 10 cm or 2° , it increases further. Beyond this, the performance of source localization will start to suffer.

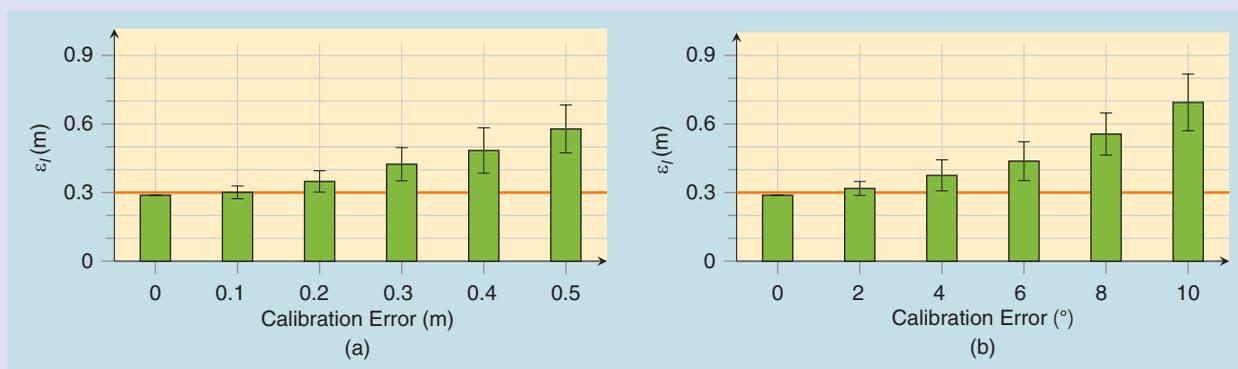


FIGURE S1. The mean speaker localization error $\varepsilon_l(m)$ as a function of (a) mean position and (b) orientation geometry calibration error.

(Figure 1). These three application scenarios, among which only the last two are relevant for WASNs, can be characterized as follows:

- 1) The first scenario (S1) addresses the calibration of individual microphones that are arranged in a microphone array of small geometric dimension. This configuration is characterized by the fact that the microphones are so close to each other that there exists some acoustic coherence between the captured signals. Furthermore, one can expect that all microphones share the same time base. The calibration of the positions of the individual microphones within the array is termed *array shape calibration*.
- 2) If each sensor node consists of a single microphone and the sensor nodes are distributed in an environment such that the microphones no longer form a compact array, the calibration process is called *microphone configuration calibration* (S2). Because of the distribution of the microphones, time synchronization between the sensor nodes cannot, in general, be expected.
- 3) The third scenario (S3) addresses the calibration of sensor nodes that are distributed in the environment, and where each sensor node consists of a small microphone

array. This task is termed *array configuration calibration*. The relative geometric arrangement of the microphones within each array is assumed to be known. Since each node is composed of more than one microphone, the arrangement of the microphones within a node is given in terms of the array position and its orientation in space. Furthermore, the absence of time synchronization between the sensor nodes has to be expected, while the microphones within each node usually share the same time base.

Approaches to geometry calibration

The common goal of all geometry calibration algorithms is the estimation of the geometric sensor arrangement. In the following, the sensor positions will be denoted by \mathbf{m}_m , $m = 1, \dots, M$, where M indicates the number of sensor nodes or microphones. The microphones are assumed to be omnidirectional, so the orientation of the individual microphones does not need to be estimated. In the case of array configuration calibration, \mathbf{m}_m refers to the center of the m th array. To be able to infer the microphone positions within the array, the orientation of the sensor node also needs to be determined. It consists of the azimuth γ_m

for two-dimensional localization and of the azimuth and elevation (γ_m, φ_m) for three-dimensional localization. The positions are gathered in a $P \times M$ matrix $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_M]$. The dimensionality of the geometric space P can be either two-dimensional ($P = 2$) or three-dimensional ($P = 3$).

Microphone position self-calibration algorithms extract measurements from the received microphone signals, which depend on the geometric arrangement of the microphones in relation to each other or to a sound source. Four basic types of acoustic measurements can be distinguished, as illustrated in Figure 2:

- 1) Pairwise distance (PD) measurements (M1) $\tilde{d}_{m,n}$ can be derived from measuring the noise coherence between microphones m and n .
- 2) Time of arrival (ToA) measurements (M2) are obtained by receiving sound from a number of positions that will, in general, be unknown. The ToA at microphone m of a sound emitted at source k at time t_k is denoted as $t_{k,m}$. The time difference $t_{k,m} - t_k$, which is also referred to as *time of flight (ToF)*, is proportional to the source-to-microphone distance in the case of a direct sound propagation from source to sensor.
- 3) Time difference of arrival (TDoA) measurements (M3) also use a number of external source positions. However, the source signals' emission times are unknown. We denote the time delay between nodes m and n of a sound emitted at position \mathbf{e}_k by $\tau_{k,(m,n)}$.
- 4) Direction of arrival (DoA) measurements (M4), sometimes referred to as *AoA*, are measured only in the third scenario (S3), where distributed microphone arrays are considered. We denote the direction at which sound k impinges on node m with a unit norm vector $\mathbf{u}_{k,m}$.

In addition to the four acoustic measurements shown in Figure 2, briefly discussed are bimodal arrangements where video cameras are employed in addition to acoustic sensors (M5). We have not, however, discussed the visual localization problem but have concentrated on acoustic localization instead, assuming that the location of the visual sensor nodes was known.

These measurements are related to the geometric arrangement: the PDs derived from noise coherence allow for a direct computation of the geometry (M1). The ToA (M2) yields the distance between source and sensor. From TDoA measurements (M3), the difference of the propagation path from a source to two receiving microphones can be inferred, and information on a

microphone array's orientation is gleaned from DoAs (M4). Furthermore, cameras on known positions are employed to resolve ambiguities and to anchor an estimated geometry in a coordinate system (M5).

Geometry calibration approaches cannot recover absolute positions. Thus, the coordinate origin is placed, without loss of generality, at the location of the first sensor node, and the second sensor is used to align the orientation. Since the calibration usually recovers only relative positions, the estimated locations exhibit an arbitrary translation and rotation with respect to the ground truth that cannot be fixed from the acoustic measurements alone. Sometimes an arbitrary reflection can occur as well. In the case of DoA-only measurements (M4), there arises an additional scale indeterminacy.

An objective function judges the deviation of the measurements from what would be expected under an assumed geometry. The form of the objective function naturally depends on the measurements used. The sought-after geometry is the one that best predicts the actual measurements. In general, the optimization problem is nonlinear and nonconvex, exhibiting multiple local minima. To avoid unfavorable local minima, researchers introduced additional information or constraints, such as low-rank matrix approximation, the assumption that the acoustic sources are in the far field, or the collocation of at least one acoustic source and sensor.

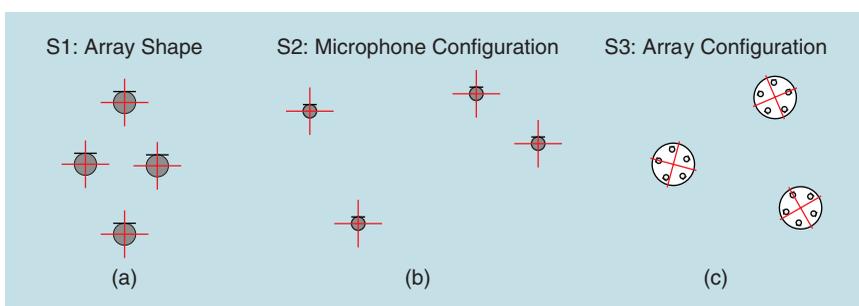


FIGURE 1. Three calibration scenarios: (a) The array shape calibration addresses the task of determining the positions of the microphones forming a compact array. (b) Microphone configuration calibration determines the position of individual microphones distributed in the room. (c) Microphone array configuration calibration seeks to determine both the position and the orientation of distributed microphone arrays whose intra-array shape is known.

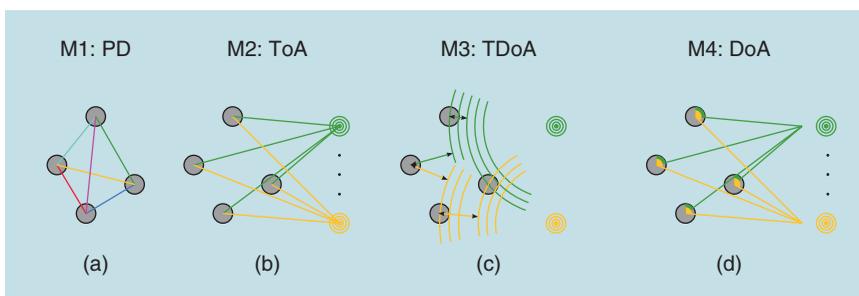


FIGURE 2. The four different types of measurements used in acoustic geometry calibration. (a) Noise coherence leads to PD measurements. (b) ToA provides distance information between a sound source and microphones. (c) The TDoA measurement yields distance-difference information between microphones in the direction of the sound source. (d) The DoA measurement provides the angle at which a sound source is observed by a sensor node.

Noise and reverberation will inevitably degrade any of the aforementioned measurements. In particular, reflections from nearby walls or objects cause errors of the measured quantities. Reverberation or the absence of a direct propagation path can especially lead to serious measurement outliers. Due to such outliers, the suitability of least squares (LS) type of objective functions, which are prevalent in geometry calibration, is somewhat questionable. However, in many cases the number of available measurements is larger than required, resulting in an overconstrained problem. Consequently, it may be attempted to remove outliers from the set of measurements. One of the most popular methods to eliminate outliers is the random sample consensus (RANSAC) [10]. Using data that provides more measurements than required, the RANSAC requires that random subsets be chosen; those that agree on a solution are used. This method is effective for removing outlier measurements and outlier geometry estimates [14], [39], [44].

To obtain the previously introduced measurements (M1–M5), three basic types of signals are used: spatially diffuse noise, dedicated point sources emitting calibration sounds such as sweeps or chirps, and natural sounds such as speech. The use of dedicated calibration signals yields the highest accuracy. However, results using speech are precise enough for most applications.

Having external sources and distributed sensor nodes raises the issue of time synchronization. On the one hand, this concerns the synchronization of the acoustic sources with the microphones and, on the other hand, the synchronization among the microphones themselves. While, in the first case, emission times will be unknown, in the second case, the measurements (M1–M4) will be affected.

Microphones in an array of small geometric dimensions are usually connected to the same sampling device, resulting in a synchronized capture of the acoustic signals. However, distributed sensor nodes have different sampling devices with different frequencies and phases. Without synchronization, the TDoA estimation of a nonmoving source would indicate a movement, since the audio signals will diverge [40]. Therefore, the question of synchronization is tightly linked to the particular formulation of the geometry calibration problem. Different assumptions on synchronization will be used to group the methods discussed. The methods involve fully synchronized setups [38], two-step approaches that first solve the synchronization and then the localization problem [11], and joint localization and synchronization methods that estimate sensor and source locations together with the time synchronization parameters [26].

In the following, an overview of approaches to geometry calibration is given, which is ordered according to the measurements (M1–M5) used. For each measurement, the optimization criterion will be formulated, and relevant publications reviewed. The notation that will be used in the following is summarized in “Notation of the Most Common Quantities.” A summary where the approaches are ordered according to the application scenario addressed concludes the survey.

Noise coherence (M1)

This class of approaches infers the microphone positions from distance measurements between pairs of microphones, which are obtained by evaluating the coherence function of a diffuse noise field. Since the assumption of a diffuse noise

Notation of the Most Common Quantities

Estimates are marked as $\hat{\cdot}$, measurements as $\tilde{\cdot}$.

To be estimated	\mathbf{m}_m	Position of sensor node m	\mathbf{M}	$P \times M$ matrix
	\mathbf{e}_k	Position of acoustic event k	\mathbf{E}	$P \times K$ matrix
	γ_m	Azimuth angle of microphone array m	γ	$1 \times M$ vector
	φ_m	Elevation angle of microphone array m	φ	$1 \times M$ vector
	$\mathbf{u}_{k,m}$	Unit length orientation vector pointing from node m toward source k in global coordinates		
Intermediate	δ_m	Absolute time offset		
	$\delta_{m,n}$	Pairwise time offset		
	t_k	Emission time of acoustic event k		
To be measured	$d_{m,n}$	Distance between nodes m and n	\mathbf{D}	$M \times M$ matrix
	$t_{k,m}$	Time of arrival of event k at node n	\mathbf{T}	$M \times K$ matrix
	$\tau_{k,(m,n)}$	Time difference of arrival of acoustic event k between nodes m and n		
	$\mathbf{v}_{k,m}$	Unit length orientation vector pointing from node m toward source k , measured in array's local coordinate system		

field is valid only for small microphone distances, this class of approaches is relevant only for the first scenario (S1). The coherence is defined as the normalized cross-power spectrum of two microphone signals. In the case of a diffuse noise field, it exhibits a sinc shape. The nulls of the sinc are proportional to the distance between the microphones. By fitting the theoretical diffuse noise coherence (DNC) function to its measurement, an estimate of the distance between the microphones can be obtained [24]. Note that the estimation of the coherence function requires the two microphones to be time synchronized.

Now, the microphone position matrix $\hat{\mathbf{M}}$ is determined, such that the estimated PDs $\|\hat{\mathbf{m}}_m - \hat{\mathbf{m}}_n\|$ are closest to the measured distances $\tilde{d}_{m,n}$:

$$\hat{\mathbf{M}} = \underset{\mathbf{M}}{\operatorname{argmin}} \sum_{m=1}^M \sum_{n=m+1}^M (\|\mathbf{m}_m - \mathbf{m}_n\| - \tilde{d}_{m,n})^2. \quad (1)$$

This optimization problem can be solved in a closed form using multidimensional scaling (MDS) [3]. Given all squared PD measurements $\tilde{d}_{m,n}^2$ that are arranged in distance matrix $\tilde{\mathbf{D}}$, MDS will find the spatial configuration of microphones. The fundamental insight is that the microphone configuration can be derived by eigenvalue decomposition from the scalar product matrix $\mathbf{B} = \mathbf{M}^T \mathbf{M}$. This matrix \mathbf{B} can be computed from the squared distance matrix, as shown in the following. Since the largest variance in this matrix is caused by the geometric displacement, the P -dimensional subspace wherein the microphones reside is spanned by the eigenvectors corresponding to the P largest eigenvalues of \mathbf{B} . Here we give only a brief overview and refer to [9] for further details.

Using the matrix of squared distance measurements $\tilde{\mathbf{D}}$, an estimate of the scalar product matrix $\hat{\mathbf{B}}$ is computed as

$$\hat{\mathbf{B}} = -\frac{1}{2} \mathbf{Q} \tilde{\mathbf{D}} \mathbf{Q}, \quad \text{where } \mathbf{Q} = \mathbf{I} - \frac{1}{M} \mathbf{1} \mathbf{1}^T \quad (2)$$

is the row and column-wise centering matrix. Here \mathbf{I} is a $M \times M$ identity matrix and $\mathbf{1}$ an m -dimensional column vector of ones. Since $\hat{\mathbf{B}}$ is symmetric and positive semidefinite, it can be decomposed into

$$\hat{\mathbf{B}} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T. \quad (3)$$

The diagonal matrix $\mathbf{\Lambda}$ contains the eigenvalues, while \mathbf{V} is the matrix composed of the corresponding unit length eigenvectors. Given that $\mathbf{B} = \mathbf{M}^T \mathbf{M}$, we see the relation to the singular value decomposition (SVD) as $\mathbf{V} \mathbf{\Sigma} \mathbf{W}^T = \operatorname{svd}(\mathbf{M})$, with $\mathbf{\Sigma} = \mathbf{\Lambda}^{(1/2)}$. The P largest eigenvalues are used to compute the geometry estimate as

$$\hat{\mathbf{M}} = \mathbf{V}_P \mathbf{\Sigma}_P, \quad (4)$$

where $\mathbf{\Sigma}_P$ is obtained from $\mathbf{\Sigma}$ by removing all except the P largest eigenvalues. Similarly, \mathbf{V}_P is the correspondingly truncated matrix of eigenvectors.

In their experiments, McCowan et al. [23] reported a microphone positioning error of around 1.5 cm. Hennecke et al. [14] performed their experiment in a reverberant conference room and achieved an accuracy of around 1 cm, with the array not too close to a wall. Otherwise, reflections from nearby walls introduce a directional bias. This shows that the assumption of diffuseness has to hold to obtain precise results.

Instead of directly matching the theoretical DNC function with the measured one, Velasco et al. [46] recently derived a model for the expected generalized cross-correlation with phase transform (GCC-PHAT) output in a diffuse noise field. With this approach, they were able to decrease the error of the PD estimates to around 0.5 cm in moderately reverberant environments.

The PDs constitute a matrix with much higher dimension than its rank. Therefore, it is possible to derive the geometry with a subset of distance measurements. Taghizadeh et al. [41] developed a low-rank matrix completion strategy to deal with incomplete measurements and to handle configurations where some microphones are not inside a compact array but half a meter away.

As an alternative to MDS, Asaei et al. [1] used nonnegative matrix factorization (NMF) to derive the microphone positions from PD measurements. The low-rank property of the distance matrix was exploited to estimate the missing values by NMF.

ToA (M2)

ToA approaches measure the arrival times of acoustic events. Assuming a point source and a direct propagation path, the ToA is proportional to the distance between the source and the microphone. Let \mathbf{e}_k , $k = 1, \dots, K$ denote the location of an acoustic event k , which will, in general, be unknown. All event locations are gathered in a $P \times K$ matrix $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_K]$. The ToA of the event k at the microphone at position \mathbf{m}_m is given by [11] as

$$t_{k,m} = \frac{\|\mathbf{m}_m - \mathbf{e}_k\|}{c} + t_k - \delta_m, \quad (5)$$

where, t_k denotes the onset time of the event, δ_m is the internal recording delay, while c refers to the speed of sound. Knowledge of t_k and δ_m assumes that source and sensor share the same time base and that the signal at the source is available. While the former requires time synchronization, the latter assumes an artificial sound source such as a loudspeaker. Relevant work is discussed where this is the case, before the more challenging case of joint localization and synchronization is examined.

Artificial sound sources and common time base

If t_k and δ_m are known, we can immediately compute the distance between \mathbf{m}_m and \mathbf{e}_k from the measured ToA $\tilde{t}_{k,m}$:

$$\tilde{d}_{k,m} = (\tilde{t}_{k,m} - t_k + \delta_m) \cdot c, \quad (6)$$

where the term in parentheses corresponds to the ToF. The ToA $\tilde{t}_{k,m}$ can be obtained by cross-correlating the microphone

signal with the calibration sound. The corresponding objective function, which jointly estimates all source and sensor positions, is

$$(\hat{\mathbf{M}}, \hat{\mathbf{E}}) = \underset{\mathbf{M}, \mathbf{E}}{\operatorname{argmin}} \sum_{m=1}^M \sum_{k=1}^K (\|\mathbf{m}_m - \mathbf{e}_k\| - \tilde{d}_{k,m})^2. \quad (7)$$

Note the similarity with (1). Thus, a direct solution of (7) is possible by employing a variant of the MDS algorithm called base point MDS (BMDS). It infers the microphone positions from PDs between the microphones and the source positions [4]. First, it computes the coordinates of a P -dimensional basis from the PDs of $P + 1$ nodes. Second, the distance measurements to the base points are used to infer the microphone positions relative to the base points. Third, a full distance matrix is constructed to run the conventional MDS algorithm.

Sachar et al. [38] used a fixed loudspeaker construction composed of four speakers on the edges of a pyramid that emitted calibration pulses. They calibrated a small 16-element microphone array with an accuracy of 0.8 cm and a large aperture array consisting of 448 microphones with an accuracy of 3 cm. Contini et al. [6] used a synchronized loudspeaker, which was moved to 25×5 positions of a rectangular grid, and white noise as a calibration signal. They were able to calibrate a linear array with an accuracy of 1 cm in both an anechoic and a strongly reverberant room.

Crocco et al. [8] used chirp signals with known emission times to obtain precise ToA estimates. An elaborate formulation of the objective function allows exploiting the constraints that the sensor and event locations are rank P matrices. By using SVD and a rank approximation technique, they were able to derive a closed-form solution in the affine space. The transformation of this solution into the Euclidean space requires an estimation of a matrix with P^2 unknown parameters. Thus, the number of unknowns became independent of the number of sensors and events. The estimation of this matrix still leads to a nonlinear optimization problem. However, this is much simpler to solve than an optimization of (7). In an experiment with



FIGURE 3. An ad hoc array composed of smartphones. (Photo used courtesy of TU Dortmund.)

eight microphones and 21 source positions, they reported an accuracy of 1 cm [7].

Active devices

Several methods use active devices such as smartphones or laptops. Here, microphones and loudspeakers are colocated. PDs can be obtained from the correlation of a calibration pulse played back by the loudspeaker at one node with the signal received by the microphone at other nodes. No time synchronization is necessary, since the emission offset is canceled out by using a pair of devices as both sender and receiver. Typically, an initial estimate is computed by MDS. Thereafter, a maximum likelihood (ML) estimation is performed, incorporating the known distance between loudspeaker and microphone in an active device [13], [35].

Raykar et al. [36] derived a joint clustering-based ML estimation. In an experiment with laptops on a table, they achieved an accuracy of around 7 cm without time synchronization and 3 cm with it. Hennecke et al. [13] used smartphones lying close to one another on a table as shown in Figure 3. Using the known speaker-microphone distance and the known smartphone orientation in the second step improved the estimate and allowed resolution of the invariance to mirrored solutions. They achieved an accuracy of 7 cm for four smartphones, which were approximately 40 cm apart from each other.

Joint localization and synchronization

If the signals' onset times and recording delays of the ToA measurements are unknown, the localization problem is considerably more difficult. Gaubitch et al. [11] suggested a two-step approach, where first the timing information and then the locations were estimated. For both estimation procedures, the low-rank structure of the sensor and source location matrices was exploited.

First (5) is rewritten as

$$\frac{\mathbf{m}_m^T \mathbf{m}_m + \mathbf{e}_k^T \mathbf{e}_k - 2\mathbf{m}_m^T \mathbf{e}_k}{c^2} = t_{k,m}^2 + t_k^2 + \delta_m^2 - 2(t_{k,m}t_k - t_{k,m}\delta_m + t_k\delta_m). \quad (8)$$

Next, the equations for $m = 1$ and for $k = 1$ are subtracted from (8). If this is done for $k = 2, \dots, K$ and $m = 2, \dots, M$, and the ToA measurements $\tilde{t}_{k,m}$ are used, the resulting system of equations can be expressed in matrix form as

$$\frac{-2\tilde{\mathbf{M}}^T \tilde{\mathbf{E}}}{c^2} = \tilde{\mathbf{T}} + \mathbf{\Gamma}(\boldsymbol{\theta}), \quad (9)$$

where $\tilde{\mathbf{M}}$ is the $P \times (M - 1)$ dimensional matrix of the microphone locations relative to the first microphone, with entry $(\mathbf{m}_m - \mathbf{m}_1)$ in the $(m - 1)$ st column, and where $\tilde{\mathbf{E}}$ is the $P \times (K - 1)$ -dimensional location matrix of the acoustic events relative to the first event, with entry $(\mathbf{e}_k - \mathbf{e}_1)$ in the $(k - 1)$ st column. Furthermore, $\tilde{\mathbf{T}}$ contains the squares of the measured ToA values, and the matrix $\mathbf{\Gamma}(\boldsymbol{\theta})$ gathers the terms that depend on the unknown timing parameters $\boldsymbol{\theta} = [t_2, \dots, t_k, \delta_1, \dots, \delta_M]$.

Since the rank of both $\bar{\mathbf{M}}$ and $\bar{\mathbf{E}}$ is P , the rank of the matrix on the left-hand side of (9) is equal to P , which is usually much smaller than both M and K . The matrix $\mathbf{\Gamma}(\boldsymbol{\theta})$ can be viewed as correcting $\tilde{\mathbf{T}}$, such that $\tilde{\mathbf{T}} + \mathbf{\Gamma}(\boldsymbol{\theta})$ also has rank P . This observation opens the way to determining the timing parameters $\boldsymbol{\theta}$ [11]: first, determine the best rank- P approximation of the right-hand side of (9). This can be achieved via SVD. Then determine the parameters $\boldsymbol{\theta}$ such that the distance between the rank- P approximation and $\tilde{\mathbf{T}} + \mathbf{\Gamma}(\boldsymbol{\theta})$ is as small as possible. These two steps are alternated until convergence. In a subsequent work, the authors employed an alternative low-rank approximation method, the structured total LS algorithm, which was much faster [15]. However, this was achieved by playing back a sequence of chips with known timing.

Once the timing parameters $\boldsymbol{\theta}$ are estimated, the location matrices $\bar{\mathbf{M}}$ and $\bar{\mathbf{E}}$ are determined, again exploiting their rank- P property. An SVD gives

$$\bar{\mathbf{M}}^T \bar{\mathbf{E}} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{W}^T, \quad (10)$$

from which $\bar{\mathbf{M}}^T$ and $\bar{\mathbf{E}}$ can be recovered as

$$\bar{\mathbf{M}}^T = \mathbf{U}_P \mathbf{C} \text{ and } \bar{\mathbf{E}} = \mathbf{C}^{-1} \boldsymbol{\Sigma}_P \mathbf{W}_P^T, \quad (11)$$

where $\boldsymbol{\Sigma}_P = \boldsymbol{\Sigma}$, all except for the P largest eigenvalues are truncated, and \mathbf{U}_P and \mathbf{W}_P consist of the corresponding left and right singular vectors, respectively.

The remaining problem is the estimation of the $P \times P$ matrix \mathbf{C} , which is much easier than the estimation of $\bar{\mathbf{M}}$ and $\bar{\mathbf{E}}$, because it is of much lower dimension. In [11], it was formulated as a nonlinear LS problem, while [8] showed that a closed-form solution can be found if one sensor is colocated with a source. If the timing parameters are known, only the second part of the algorithm needs to be carried out, as described in [8].

Using microphones randomly distributed on a table, an accuracy of 2 cm was achieved with this method [11]. The TDoA measurements were obtained from recordings of handclaps. However, the authors needed to label the largest peak in each handclap signal manually to obtain precise estimates.

TDoA (M3)

The TDoA is proportional to the distance difference of a pair of microphones to the source, when the direct path from the source to the microphones exists. The TDoA itself may be measured by the maximum in the correlation of the two microphone signals, or by onset detection. To derive the geometry, this measurement is related to the positioning as follows: The TDoA from the k th source to the m th and n th sensor is given by

$$\tau_{k,(m,n)} = t_{k,m} - t_{k,n} = \frac{\|\mathbf{m}_m - \mathbf{e}_k\| - \|\mathbf{m}_n - \mathbf{e}_k\|}{c} - \delta_m + \delta_n. \quad (12)$$

Note that the onset time t_k cancels out. Here, $-\delta_m + \delta_n$ is the time offset between the recording devices.

Relevant work is discussed first, which assumes time synchronization and thus absence or, equivalently, knowledge of the delays, before we turn to the case where the delays have to be estimated as ancillary parameters.

Synchronized microphones

In reverberant environments, the steered response power with phase transform is often employed for TDoA-based localization. This is equivalent to a filter-and-sum beamformer, when steering the beamformer to all possible locations and selecting the position where the output energy is maximized [5]. The individual source positions obtained from several distributed microphone arrays located at the ceiling of a highly reverberant conference room were used by Hennenke et al. to perform a coordinate mapping [14]. In their experiments, speech and noise emitted from random positions was used together with a RANSAC scheme. They achieved an accuracy of 10 cm with speech and white noise in most cases.

If a sensor node consists of a microphone array, an acoustic camera can be formed by a delay-and-sum beamformer applied to the received signals of each array. Redondi et al. [37] used pure sinusoids as source signals to obtain acoustic images. This enables the application of computer vision techniques. They used camera models to extract positions in Cartesian coordinates, which were used as input for a subsequent coordinate mapping approach. The coordinate mapping approach recovered the rotation and translation for each microphone array to a selected reference array.

Thrun [43] showed that the localization problem can be significantly simplified if the sources are in the far field of the microphones. Then a source signal impinges on all sensors from the same angle, and the actual position of the source is immaterial. Thus, we can write $\|\mathbf{m}_m - \mathbf{e}_k\| = \mathbf{u}_k^T (\mathbf{m}_m - \mathbf{e}_k)$, where the unit-length direction vector \mathbf{u}_k only depends on the source and is independent of which microphone is considered. Then the right-hand side of (12) simplifies to $\mathbf{u}_k^T (\mathbf{m}_m - \mathbf{m}_n)/c$, which has to be compared to the measured TDoA $\tilde{\tau}_{k,(m,1)}$, leading to the overdetermined system of equations

$$\mathbf{U}^T \bar{\mathbf{M}}/c = \boldsymbol{\tau}, \quad (13)$$

where \mathbf{U} is the $P \times K$ matrix, with \mathbf{u}_k on the k th column, and $\boldsymbol{\tau}$ is the $K \times (M-1)$ matrix of measured TDoAs. Note that Thrun assumes that the δ terms are known and thus can be set to zero.

Again, the rank argument can be invoked. The rank of the left-hand side of (13) is P , and so must be the rank of the right-hand side. Thus, we can apply the same rank approximation by SVD as explained in (10) and (11) to infer the microphone positions. This approach of Thrun and others has become known as *affine structure from sound (ASfS)*.

Unsynchronized microphones

In the general case of unsynchronized microphones, the recording delays δ_m , $m = 1, \dots, M$ are different and unknown.

A joint estimation with the locations leads to the nonlinear LS problem [25]

$$\hat{\mathbf{M}}, \hat{\mathbf{E}}, \hat{\delta} = \underset{\mathbf{M}, \mathbf{E}, \delta}{\operatorname{argmin}} \sum_{m=2}^M \sum_{k=1}^K \left(\frac{\|\mathbf{m}_m - \mathbf{e}_k\| - \|\mathbf{m}_1 - \mathbf{e}_k\|}{c} - \delta_m + \delta_1 - \tilde{\tau}_{k,(m,1)} \right)^2 \quad (14)$$

In [25], the optimization problem was solved using an auxiliary function-based algorithm, which is an extension of the expectation-maximization (EM) algorithm. The alternating optimization of the positions and timing parameters was shown to have better convergence properties than gradient descent.

As an alternative to this one-stage scheme that estimates the unknown positions and timing parameters simultaneously, Wang et al. [47] recently proposed a two-stage method. This method again uses a rank approximation technique. While the aforementioned rank approximation-based algorithms [8], [11], [12] are suitable for ToA and require at least the onset times or the internal delays to be known, Thrun [43] was able to work with TDoA; however, both onset time and internal delays needed to be known. Wang et al. [47] coped with these limitations and derived an approach that uses TDoA measurements and estimates the unknown onset times and internal delays. Furthermore, Wang et al. showed that the combination of their algorithm and the EM algorithm-based calibration from [25] can outperform the individual methods.

An interesting insight is that the maximum TDoA (mTDoA) can be used to estimate PDs [28]. The maximum of the TDoA values is obtained if the source is aligned with the microphone-to-microphone direction (the so-called endfire position). To achieve this condition, sound from a variety of distributed source positions is used and the maximum is computed over the whole period over all of them. Unlike the noise coherence method, this approach does not require a time synchronization between the microphones and is able to handle larger microphone distances. On the other hand, the intermicrophone distance will be underestimated by the mTDoA approach if no event is observed in the endfire position. Let

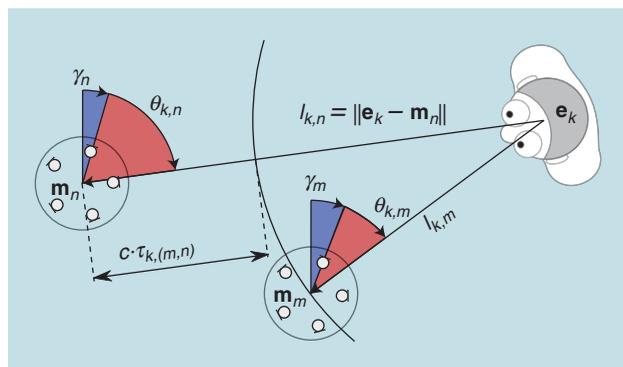


FIGURE 4. Two sensor nodes consisting of five-element circular arrays at \mathbf{m}_m and \mathbf{m}_n and a source at \mathbf{e}_k . The DoA θ is offset by the angle γ with respect to the global coordinate system. The TDoA $\tau_{k,(m,n)}$ corresponds to the difference between the distances of each array to the speaker. The illustration is in 2-D, so $\mathbf{v}_{k,m} = (\cos \theta_{k,m}, \sin \theta_{k,m})^T$.

$$\tilde{\tau}_{n,m}^{\max} = \max_k \{ \tilde{\tau}_{k,(m,n)} \} = \|\mathbf{m}_n - \mathbf{m}_m\|/c + \delta_m - \delta_n \quad (15)$$

and

$$\tilde{\tau}_{n,m}^{\min} = \min_k \{ \tilde{\tau}_{k,(m,n)} \} = -\|\mathbf{m}_n - \mathbf{m}_m\|/c + \delta_m - \delta_n, \quad (16)$$

where we assume that an acoustic event exists in either endfire position. Thus, it follows that the time offset can be computed as in [27], i.e., that

$$\hat{\delta}_{m,n} = \delta_m - \delta_n = \frac{1}{2} (\tilde{\tau}_{m,n}^{\max} + \tilde{\tau}_{m,n}^{\min}), \quad (17)$$

and that the distance is

$$\hat{d}_{m,n} = (c/2) (\tilde{\tau}_{m,n}^{\max} - \tilde{\tau}_{m,n}^{\min}). \quad (18)$$

Now, MDS can be invoked again to derive an estimate of the microphone geometry [28]. The performance of the result depends on whether the mTDoA is observed. However, reasonable results are obtained even if this is not true for all pairs. Both Parviainen et al. [26] and Pertilä et al. [28] performed experiments with smartphones in a meeting room, where they achieved around a 12-cm accuracy.

DoA (M4)

DoA-based calibration can be conducted only for the third scenario (S3) described in the “Application Scenarios” section, since it requires a microphone array per sensor node rather than a single microphone to acquire DoA estimates.

An objective function is formed by comparing the direction of the k th acoustic event, as measured by array m , $\tilde{\mathbf{u}}_{k,m}$, with the direction predicted by the assumed geometry: $\mathbf{u}_{k,m} = (\mathbf{m}_m - \mathbf{e}_k) / \|\mathbf{m}_m - \mathbf{e}_k\|$. Since we are working with directions, a Euclidean distance measure is inappropriate. A cosine distance measure is used instead

$$1 - \cos(\angle(\tilde{\mathbf{u}}_{k,m}, \hat{\mathbf{u}}_{k,m})) = 1 - \tilde{\mathbf{u}}_{k,m}^T \mathbf{u}_{k,m}. \quad (19)$$

The vector $\tilde{\mathbf{u}}_{k,m}$ describes the measurement of the impinging angle with respect to a global coordinate system. However, a measurement $\tilde{\mathbf{v}}_{k,m}$, which can be obtained from the microphone signals, is located in the local coordinate system of the sensor node. For an illustration in two dimensions, see Figure 4. Since the array exhibits an unknown azimuth γ_m and elevation φ_m with respect to the global coordinate system, the rotation has to be compensated. This is achieved using a rotation matrix $\mathbf{R}(\gamma_m, \varphi_m)$

$$\tilde{\mathbf{u}}_{k,m} = \mathbf{R}^{-1}(\gamma_m, \varphi_m) \tilde{\mathbf{v}}_{k,m}. \quad (20)$$

This transformation allows the combination of all measurements in a common objective function, if (20) is plugged in to (19) and we eventually arrive at

$$\hat{\mathbf{M}}, \hat{\mathbf{E}}, \hat{\gamma}, \hat{\varphi} = \underset{\mathbf{M}, \mathbf{E}, \gamma, \varphi}{\operatorname{argmin}} \sum_{m=1}^M \sum_{k=1}^K (1 - \tilde{\mathbf{v}}_{k,m}^T \mathbf{R}(\gamma_m, \varphi_m) \mathbf{u}_{k,m}). \quad (21)$$

The optimization of (21) was carried out in [19] by using the Newton algorithm. Since [19] was solely working on DoA estimates, neither a time synchronization between the sensor nodes nor a synchronization between the source and sensors was required. The only requirement was that the microphones within an array were synchronized, since the DoA is eventually measured by TDoAs. To handle noisy measurements, the RANSAC framework was employed.

A DoA-based localization is obviously unable to determine the scale of the geometry. Scale ambiguity can be resolved by employing an additional TDoA measurement, as proposed in [39]. These measurements provide distance differences, which are used to scale the DoA-only calibration result, such that the geometry matches the distance differences. The overall system is shown in Figure 5. However, the interarray TDoA estimation requires a time synchronization between sensor nodes. Jacob et al. [16] showed that the knowledge of the intra-array geometry of a circular array is sufficient to solve the scale ambiguity problem. Thus, the TDoA can be omitted.

The method introduced in [31] combined DoA measurements from the individual microphone arrays and TDoA measurements between the arrays. For each pair m, n of arrays, the position and orientation have to fulfill the geometric relations with respect to the measured TDoA and DoA, as illustrated in Figure 4. Given an estimate of the azimuthal $\hat{\gamma}_m, \hat{\gamma}_n$ and elevation displacements $\hat{\phi}_m, \hat{\phi}_n$, and estimates of the positions $\hat{\mathbf{m}}_m, \hat{\mathbf{m}}_n$, the measured directions

$\tilde{\mathbf{u}}_{k,m}$ and $\tilde{\mathbf{u}}_{k,n}$ can be used to compute the source position by triangulation [31]

$$\begin{aligned} \hat{\mathbf{e}}_k(m, n) &= \hat{\mathbf{m}}_m + \hat{l}_{k,m} \tilde{\mathbf{v}}_{k,m} \mathbf{R}^{-1}(\hat{\gamma}_m, \hat{\phi}_m) \\ &= \hat{\mathbf{m}}_n + \hat{l}_{k,n} \tilde{\mathbf{v}}_{k,n} \mathbf{R}^{-1}(\hat{\gamma}_n, \hat{\phi}_n). \end{aligned} \quad (22)$$

Here, the notation $\hat{\mathbf{e}}_k(m, n)$ indicates that the source position estimate is obtained from measurements of arrays m and n . The line intersection provides estimates $\hat{l}_{k,m} = \|\mathbf{e}_k - \mathbf{m}_m\|$ and $\hat{l}_{k,n}$ of the distances to the source. If either of these distances is negative, there is no intersection and the solution points to a mirrored position.

Next, the measured TDoA is compared with the value predicted by the current estimates of source and sensor positions. As can be seen in Figure 4, they are related as $c \cdot \tau_{k,(m,n)} = l_{k,n} - l_{k,m}$. This comparison allows the formation of an objective function that is determined for three or more source positions. It is used to estimate the geometry hierarchically (Figure 6). First, the geometry of all pairs $(1, m), m = 2, \dots, M$ is estimated individually. Then these estimates are used as a starting point to estimate the geometry of all microphone arrays jointly. The procedure is repeated to improve the estimation and remove a bias that can be the result of an individual DoA error. Random subsets of three or more positions are used, and the final geometry estimation result is obtained as the average of several such subsets.

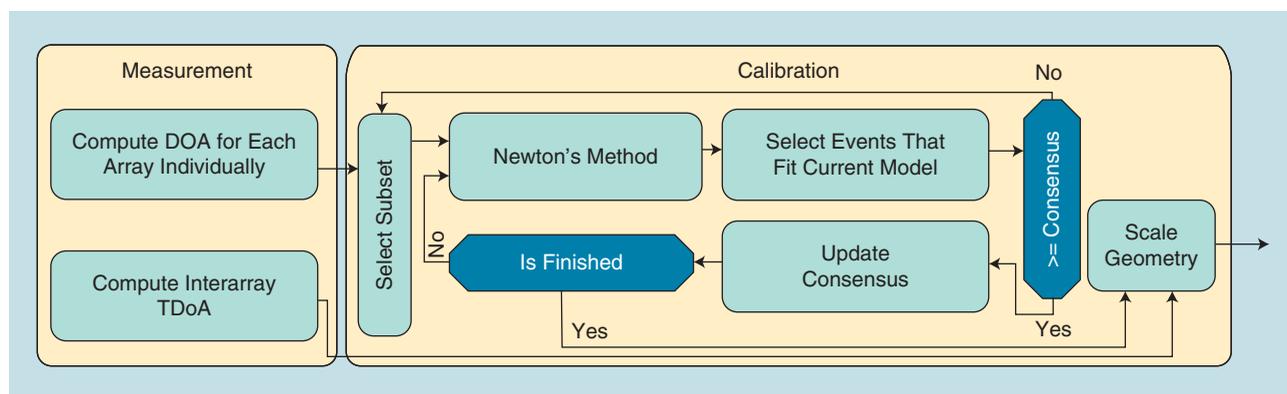


FIGURE 5. A DoA-based calibration embedded into a RANSAC framework [19] and combined with scale factor estimation from [39].

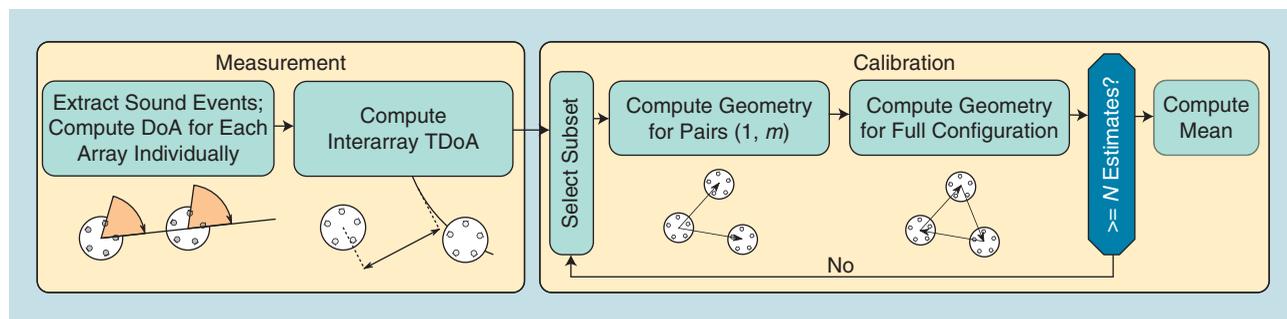


FIGURE 6. A DoA-TDoA method for the calibration of distributed microphone arrays [31].

The combined DoA-TDoA method for microphone array configuration calibration achieved a position accuracy of 10 cm and an angular accuracy of 2° using speech, compared to 1 cm and 1° accuracy when using noise [31]. These values were measured in a 2-D scenario with table-top microphones.

Visual support (M5)

Acoustic sensors are often accompanied by visual ones. This allows joint calibration of the audiovisual sensor network employing both the acoustic and visual modality. However, the discussion of the geometry estimation of a camera network and thus the joint calibration of audiovisual sensor networks is beyond the scope of this survey. We will only briefly discuss how known positions of cameras can be used to resolve the translation and rotation indeterminacy and, if required, also the scale indeterminacy of a purely acoustic geometry calibration.

If a speaker is tracked separately, both by a camera network with known camera positions and by a microphone array whose positions have been estimated by one of the aforementioned geometry calibration methods, the microphone positions have to be embedded into the coordinate system provided by the camera network. This is achieved by matching the acoustic and visual speaker trajectory [17], thus fixing the translation, rotation, and scale of the acoustic sensor network. An alternative to this postmatching of trajectories is “online” joint audiovisual localization, as described in [18].

In [30], Plinge and Fink assumed that the visual modality provides the absolute positions of the speakers, from which the microphone geometry is calibrated using DoA measurements. This approach can be seen as a reverse localization problem, where the source positions are known and the sensor positions need to be estimated. As the relation of the DoA and the source position is straightforward (Figure 4), each node’s position and orientation can be estimated along with the distances $\mathbf{l}_m = (l_{1,m} \dots l_{K,m})^T$. For any set of three or more speaker positions \mathbf{e}_k , the corresponding equations can be combined into an overdetermined system of equations. The geometric arrangement of the microphone arrays can be determined by solving the resulting objective function

$$\hat{\gamma}_m, \hat{\mathbf{m}}_m, \hat{\mathbf{l}}_m = \underset{\mathbf{m}_m, \gamma_m, \mathbf{l}_m}{\operatorname{argmin}} \sum_k \|\mathbf{e}_k - \mathbf{m}_m - l_{k,m} \tilde{\mathbf{v}}_{k,m} \mathbf{R}^{-1}(\hat{\gamma}_m, \hat{\phi}_m)\|. \quad (23)$$

Given reasonably accurate visual localizations, this objective function is approximately convex, since it increases monotonously with both position and orientation errors. Therefore, gradient descent can be used, and a solution is found regardless of the initialization.

By embedding the microphone positions into the coordinate system of the camera system, the cross-modality speaker tracking and localization outperformed a solely acoustic and a solely visual localization. Thus, it is beneficial to embed the microphone positions into the coordinate system of the camera system.

Summary

Table 1 gives an overview of the calibration approaches. They are grouped by the application scenarios (S1–S3) introduced in the “Application Scenarios” section and the measurements employed (M1–M5). The references given are accompanied by a few remarks concerning the methodology and types of experiments conducted. In the following, we summarize some design considerations.

Objective functions

MDS or BMDS is the preferred approach if PDs are given. It is used in all three application scenarios. If solely compact arrays are considered, the distance can be obtained by the diffuse noise approach. For distributed active devices, the required distance measurements can be obtained by correlating transmitted and received signals. However, the mTDoA approach is able to work with ambient sounds or speech signals. ToA, TDoA, and DoA lead to nonlinear LS problems, where care has to be taken to avoid unfavorable local minima.

Number of source positions

The required number of source positions for a successful calibration varies significantly. Only the diffuse noise approach does not require any active sources—except for the presence of diffuse noise. The number of different source positions used in the other experiments ranges from about ten to more than 100, depending on the method and objective function used. In the past, several approaches used loudspeakers mounted at fixed positions [20], [38]. With the advent of more elaborate methods, a single moving source could be utilized. A particularly convenient method is to employ a handheld smartphone playing noise or chirps [12], [31].

Reverberation will degrade any measurement. Apart from increasing the number of source positions, additional steps to increase the robustness have to be added. These include robust estimation [29] or RANSAC [39], [45].

Signal types

We can distinguish three classes of sounds used: diffuse noise, dedicated calibration signals, and natural sounds such as speech. Spatially diffuse noise can be used only for the calibration of small microphone arrays [3], [41].

Many calibration methods use calibration signals with good correlation properties, such as white noise or chirps, that allow for exact TDoA measurements. The use of a calibration signal provides more accurate calibration results than the use of natural sounds.

For online and ad hoc scenarios, speech is preferred—but the use of speech is challenging, since it will provide less accurate measurements compared to white noise due to its less sharp autocorrelation function. Additionally, when speech is used, it is necessary to preprocess the microphone signal by voice activity detection or signal classification to exclude badly localized sounds produced, for example, by furniture, chairs, footsteps, or doors [33]. The accuracy of speech-based

Table 1. An overview of the calibration methods discussed, ordered by scenario and measurement.

	Array Shape (S1)	Microphone Configuration (S2)	Array Configuration (S3)
PD (M1)	<ul style="list-style-type: none"> • MDS <ul style="list-style-type: none"> – Manual measurements [3] – Diffuse noise [24], [46] • Rank approximation <ul style="list-style-type: none"> – Diffuse noise + far-field microphones [41] 	<ul style="list-style-type: none"> • NMF <ul style="list-style-type: none"> – Incomplete + Noisy distance [1] 	
ToA (M2)	<ul style="list-style-type: none"> • BMDS <ul style="list-style-type: none"> – Known distances [4] • Direct minimization <ul style="list-style-type: none"> – Known loudspeaker configuration [6], [38] 	<ul style="list-style-type: none"> • Direct minimization <ul style="list-style-type: none"> – Known loudspeaker configuration [38] • Rank approximation via SVD <ul style="list-style-type: none"> – Known emission time [7], [8] – Impulse train [12] – Manually labeled handclaps [11] • Active devices <ul style="list-style-type: none"> – One microphone and event are colocated [35] – Unsynchronized + MDS [13], [36] 	
TDoA (M3)		<ul style="list-style-type: none"> • Direct minimization <ul style="list-style-type: none"> – No initialization [34] – Auxiliary function [25] • Rank approximation via SVD <ul style="list-style-type: none"> – Far-field sources [43] – Unsynchronized [47] • mTDoA <ul style="list-style-type: none"> – MDS + speech [26]–[28] 	<ul style="list-style-type: none"> • Coordinate mapping <ul style="list-style-type: none"> – Direct minimization [45] – Random sampling [14] – Acoustic camera [37]
DoA (M4)			<ul style="list-style-type: none"> • RANSAC <ul style="list-style-type: none"> – Random walk + speech [16], [19] • Intra-array TDoA <ul style="list-style-type: none"> – Fixed speaker positions [32] – Random walk [39]
Visual support (M5)			<ul style="list-style-type: none"> • Audiovisual <ul style="list-style-type: none"> – Trajectory mapping [17] – Visual speaker localization [30] – Joint calibration [18]

methods tends to be lower, but fortunately high enough for practical applications.

Synchronization

If the clocks of transmitter and receiver are synchronized, PDs can be obtained from ToA measurements. TDoA measurements require only a synchronization among the microphones. A further relaxation is possible if only DoA measurements are incorporated. Some algorithms couple the estimation of the sampling deviation and the calibration process itself [11], [25], [47], while the mTDoA method elegantly removes a potential unknown delay [26], [27] (cf. the treatment of timing difference by the mTDOA approach in the “Unsynchronized Microphones” section).

Experimental evaluation of selected methods

The section “Approaches to Geometry Calibration” provided an overview about a broad range of geometry

calibration algorithms. The authors of the corresponding publications evaluate their algorithms usually on their proprietary data sets, which makes a comparison among different approaches difficult. This section tries to fill this gap and provides a comparison under a common evaluation framework. We conducted experiments for all three application scenarios, and for each scenario we evaluated a selection of algorithms in a two-dimensional calibration experiment in a reverberant laboratory environment. We also did our best to correctly implement and optimize those algorithms that we have not proposed. We do not claim, however, that we achieved their best possible performance. Table 2 provides an overview of the algorithms selected and the scenarios where they have been applied. In the following, we first describe the test environment and the performance evaluation metrics used. Afterward, we present the results for each of the three scenarios. The evaluation is concluded by a short summary.

Table 2. An overview of the methods used in the evaluation.

Array Shape (S1)	Microphone Configuration (S2)	Array Configuration (S3)
DNC + MDS [24]	ToA rank [11]	DoA + TDoA scaling [19], [39]
mTDoA + MDS [28]	mTDoA + MDS [28]	DoA-TDoA [31]
ASfS [43]	ASfS [43]	DoA + Video [30]

Evaluation setups and metrics

The location was a highly reverberant $3.7 \text{ m} \times 6.8 \text{ m} \times 2.6 \text{ m}$ conference room of a smart house installation at TU Dortmund University. Signals from three circular microphone arrays that were arranged in an irregular triangle of an approximate edge size of 1 m were recorded at 48 kHz (Figure 7). Each array was embedded in the table and consisted of five microphones arranged equidistantly on a circle of radius 5 cm. The signals were captured synchronized. A reverberation time (T_{60}) of 0.67 seconds was calculated using a blind estimation algorithm [21]. Five cameras mounted at the ceiling captured the scene at 10 frames/second and 384×288 pixel resolution. They have a field of view of $48^\circ \times 36^\circ$. Acoustic events were produced from ten locations around the table. For the first recording, a smartphone

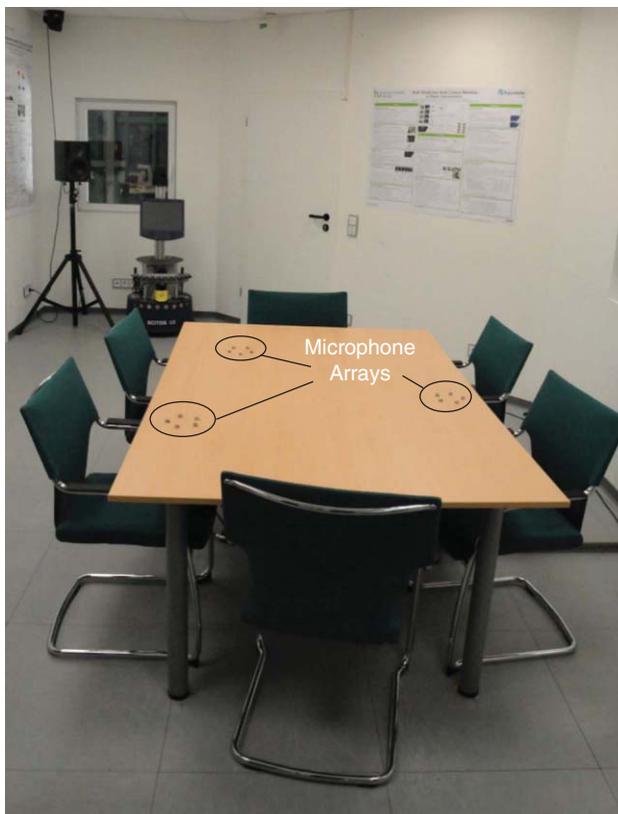


FIGURE 7. The recording setup. The conference table is embedded with three circular microphone arrays.

was held at the same height as the microphones, and a white noise signal was played back. In the second case, a speaker was either sitting (four positions) or standing (six positions) at the table. Consequently, his mouth was approximately 0.4 m or 0.7 m above the microphones, respectively.

As mentioned previously, the estimated geometry $\hat{\mathbf{M}}$ exhibits an arbitrary translation and rotation with respect to the true geometry that needs to be removed before an error can be measured. To this end, a translation vector \mathbf{t} and a rotation matrix \mathbf{R} are determined by SVD, such that the mean location error between the estimated microphone positions after correction and the true microphone positions is minimized. Therefore the estimated positions after correction are given by $\hat{\mathbf{m}}_i = \mathbf{R}\hat{\mathbf{m}}_i + \mathbf{t}$, where $\hat{\mathbf{m}}_i$ represents the original estimates. The performance measure for the positions is the root-mean-square (RMS) error

$$\epsilon_p = \sqrt{\frac{1}{M} \sum_{i=1}^M \|\mathbf{m}_i - \hat{\mathbf{m}}_i\|^2}. \quad (24)$$

In the case of array configuration calibration, the orientation of the arrays is also an important parameter to estimate. The estimate has an arbitrary rotation relative to the ground truth. To compensate for this, an angle d_γ is determined such that the deviation of the ground truth from the estimated angles $\hat{\gamma}_j$ after rotation by d_γ is minimal. Then the average orientation error is computed as

$$\epsilon_\gamma = \frac{1}{M} \sum_{m=1}^M |\gamma_m - d_\gamma - \hat{\gamma}_m|. \quad (25)$$

Since several algorithms solve nonlinear problems and use random initializations, each experiment is repeated ten times. The average and standard deviation computed over these runs is reported.

Microphone array shape calibration (S1)

We compare the DNC approach [24] and the mTDoA [28] algorithm, which both estimate PDs, from which the overall geometry is inferred by MDS. The comparison includes the ASfS approach [43], which originally was developed for microphone configuration calibration but can also be employed for microphone array shape calibration.

The methods achieved a positioning accuracy of 0.3–1.5 cm for the microphones of the circular array described above. The overall results for the DNC and mTDoA method were rather similar, while ASfS, which was developed for microphone configuration calibration problems, performed slightly worse (Figure 8). We observed, though, that the mTDoA + MDS method does not degrade as quickly as the DNC approach when the interelement distance is increased. The DNC approach relies on the presence of ambient diffuse noise, while the mTDoA algorithm relies in the presence of sound sources being in the endfire position of microphone pairs. Either requirement may not always be fulfilled.

Microphone configuration calibration (S2)

The microphone configuration calibration performance was evaluated on all 15 microphones of the three circular arrays. The mTDoA approach combined with MDS [28] worked for pairs of microphones from different arrays with a distance of up to about 1 m. We therefore included it in our evaluation and compared it to the ASfS approach [43] and the ToA rank approximation scheme [11].

Figure 9 compares the calibration error of the methods using either white noise or speech as calibration signals. The best localization is achieved with mTDoA, resulting in an RMS error of $4.5 \text{ cm} \pm 1.8 \text{ cm}$ using a speech signal, while $1.3 \text{ cm} \pm 0.7 \text{ cm}$ was achieved with noise excitation. The ASfS approach performed slightly worse with $6.0 \text{ cm} \pm 2.7 \text{ cm}$ and $1.8 \text{ cm} \pm 0.8 \text{ cm}$, respectively. Our implementation of the ToA rank scheme did not perform well. This might be a consequence of the arrangement, since experiments with uniformly distributed microphones performed significantly better.

Array configuration calibration (S3)

For the array configuration calibration, the DoA-TDoA method [31], the DoA + Video method [30], and the DoA + TDoA scaling approach [19], [39] were used.

The required DoAs were estimated by a neurobiologically inspired method [29], since it is robust to noise and excludes nonspeech sounds. The event locations were automatically identified as segments with a small angular variance. The error in the DoA angle estimation was around 3° . The TDoA information was extracted by computation of the GCC-PHAT.

The TDoAs over all microphone pairs had an error of around 6 cm. For the multimodal method, visual localization by background subtraction and an upper-body detector was used [30]. Seven localizations with an accuracy of 20 cm were derived for the ten detected speech segments. For the noise sequence, the ground truth positions marked on the floor where the sounds were produced were used.

The calibration results are shown in Figure 10. All methods achieved an average position error ϵ_p of less than 7 cm. The maximum position error for the DoA-TDoA method was 0.6 cm for noise and 6.0 cm for speech. For the DoA + Video approach, it was 5 cm for noise and 7 cm for speech. For the DoA + TDoA scaling method, it was 4 cm for noise and 2.5 cm for speech. The angular error is close to 1° for all methods, except for the DoA-TDoA method using speech, where only 3° was achieved.

Summary

In our experiments, array calibration could be performed quite accurately using diffuse noise or mTDoA from multiple distributed speech events. Below 1-cm precision is close to the requirements for beamforming. While the diffuse noise approach is limited to small array sizes, the latter method also allowed the calibration of distributed microphones on a table using speech or noise with 10-cm and 3-cm precision,

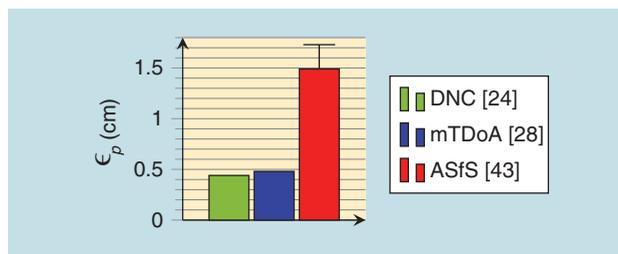


FIGURE 8. Array shape calibration. The mean position error for tabletop microphones in a reverberant smart room.

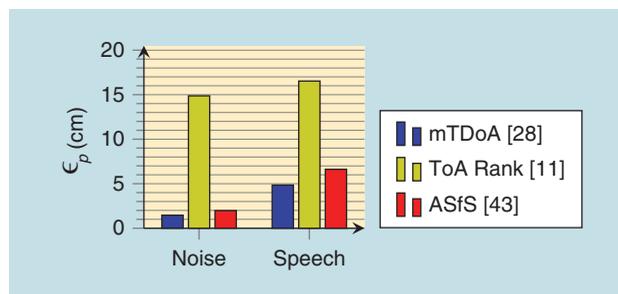


FIGURE 9. Microphone configuration calibration. The mean position error using either noise or speech as input signals.

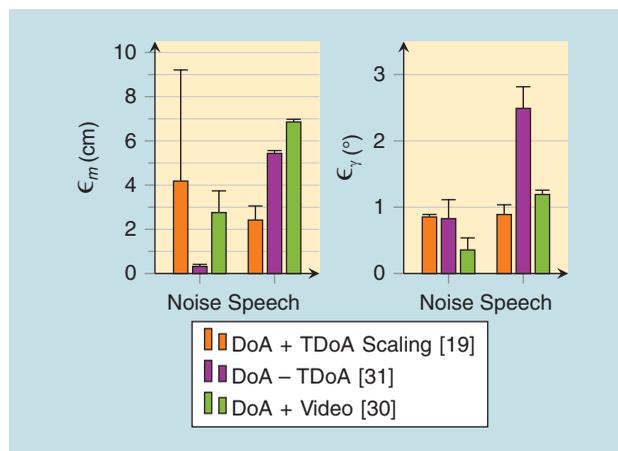


FIGURE 10. Array configuration calibration. The mean position and orientation error for three DOA-based algorithms.

respectively. This can also be achieved using the ASfS method [43], with a slightly higher error. We were able to show that state-of-the-art methods [19], [31] are capable of calibrating array configurations with an orientation error of well below 5° and a position error well below 10 cm. This provides sufficient accuracy for triangulation-based processing algorithms.

Conclusions and outlook

This article provided a survey of acoustic geometry calibration algorithms, which attempt to reveal the position of microphones solely from the acoustic signals received by them. The algorithms can be categorized on the basis of such things as the kind of acoustic signals used, the kind of position-related measurements employed, the kind of objective function used, and the necessity of synchronization.

We have chosen to organize the presentation according to two criteria. The first is based on the scenario addressed: the position estimation of individual microphones within a microphone array (S1), the localization of distributed microphones (S2), and the positioning of distributed microphone arrays (S3). The second criterion is the measurements used, from which position-related information is extracted: the noise coherence function (M1), which is related to the distance between two microphones; the ToA (M2), which is related to the distance between a sound source and a sensor; the TDoA (M3) between the signals at two microphones, which is related to the distance between the microphones and the angle at which the sound is observed; and finally the DoA (M4), from which the relative geometry of sources and sensors can be revealed. These two categorizations have been chosen to provide, on the one hand, an application-oriented point of view and, on the other hand, a technology-oriented perspective. While the first allows a practitioner to quickly identify which approach is suitable for which application, the second may help researchers to sort the different approaches according to which input data are used.

The survey and the experimental evaluation showed that the highest accuracy is achieved if dedicated calibration signals are used, such as chirps from known positions, and if transmitter and receiver are synchronized. A positioning accuracy on the order of 1 cm has been reported in all scenarios. For usability reasons, however, a localization from natural sounds such as speech is preferable. But even with speech, accuracies in the range of 5 cm are achievable, which is high enough for applications like speaker tracking.

Further improvements in accuracy and usability can be obtained along different lines of research. First, the quality of the measurements should be improved. The aforementioned measurements are already the result of some signal processing. They are often obtained from cross-correlations of acoustic signals, and these correlations are heavily affected by noise and reverberation. More robust correlation results would immediately lead to improved input data for the calibration process and thus improved calibration results.

Second, the objective functions are often nonlinear LS-type functions, which exhibit multiple local minima and which are often optimized iteratively. However, due to reverberation and reflections, the measurement error is not normally distributed, putting the suitability of LS into question. Today, mostly ad hoc countermeasures such as RANSAC are used, while a more principled approach to handle outliers has yet to come.

Furthermore, the optimization should be guided by a priori knowledge or sensible assumptions to avoid unfavorable local minima. Examples that have proven effective are the exploitation of the low-rank property of position and distance matrices or the far-field assumption. Such constraints are the more important the more complex the optimization problem becomes. A particularly complex case is the calibration of unsynchronized microphones from ToA or TDoA measurements using speech input in noisy reverberant environments. This requires the synchronization problem to be solved

before or along with the localization problem. We expect more work in this field, both because of the difficulty, and thus attractiveness, of the problem, and because of its practical importance.

Acknowledgment

This work was supported by the German Research Foundation under contracts Fi 799/5-1 and Ha3455/7-2.

Authors

Axel Plinge (axel.plinge@tu-dortmund.de) received his diploma degree with distinction in computer science with a minor in philosophy from the Technical University of Dortmund, Germany, where he worked at the Leibniz Research Centre for Working Environment and Human Factors in different areas of psychophysical research from hearing to color vision and depth perception. In two European research projects, he developed innovative speech enhancement and replacement methods for persons with severely sensory-impaired hearing. He has published several papers on a variety of topics, including human hearing and vision, speech enhancement, pattern recognition, speaker tracking, and geometry calibration.

Florian Jacob (jacob@nt.uni-paderborn.de) received his Dipl.-Ing. degree in computer engineering with a specialization in electrical engineering from Paderborn University, Germany, in 2011. He has been a research staff member in the Department of Communications Engineering since 2011. His research interests include acoustic signal processing and unsupervised geometry calibration of distributed sensor networks. He is currently working toward his Ph.D. degree.

Reinhold Haeb-Umbach (haeb@nt.uni-paderborn.de) received his Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from Rheinisch-Westfälische Technische Hochschule Aachen, Germany, in 1983 and 1988, respectively. From 1988 to 1989, he was a postdoctoral fellow at the IBM Almaden Research Center, San Jose, California, and conducted research on coding and signal processing for recording channels. From 1990 to 2001, he was with Philips Research and worked on various aspects of automatic speech recognition. Since 2001, he has been a professor of communications engineering at Paderborn University, Germany. His main research interests are statistical speech signal processing and recognition. He has published more than 150 papers in peer-reviewed journals and conferences.

Gernot A. Fink (Gernot.Fink@udo.edu) received his diploma in computer science from the University of Erlangen-Nuremberg, Germany, in 1991. From 1991 to 2005, he was with the Applied Computer Science Group at Bielefeld University, Germany, where he received his Ph.D. degree (Dr.-Ing.) in 1995 and his *venia legendi* (Habilitation) in 2002. Since 2005, he has been a professor at the Technical University of Dortmund, Germany, where he heads the Pattern Recognition in Embedded Systems Group. His research interests are machine perception, statistical pattern recognition, and document analysis. He has published more than 150 papers and a textbook on Markov models for pattern recognition.

References

- [1] A. Asaei, N. Mohammadiha, M. J. Taghizadeh, S. Doclo, and H. Bourlard, "On application of non-negative matrix factorization for ad hoc microphone array calibration from incomplete noisy distances," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Brisbane, Queensland, Australia, Apr. 2015, pp. 2694–2698.
- [2] A. Bertrand, "Applications and trends in wireless acoustic sensor networks: A signal processing perspective," in *IEEE Symp. Commun. and Vehicular Technol. in the Benelux*, Ghent, Belgium, Nov. 2011, pp. 1–6.
- [3] S. T. Birchfield, "Geometric microphone array calibration by multidimensional scaling," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Hong Kong, Hong Kong, Apr. 2003, pp. 157–160.
- [4] S. T. Birchfield and A. Subramanya, "Microphone array position calibration by basis-point classical multidimensional scaling," *IEEE Trans. Speech and Audio Process.*, vol. 13, no. 5, pp. 1025–1034, Sept. 2005.
- [5] M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*, 1st ed. Springer Science and Business Media, Berlin, Germany: Springer-Verlag, 2001.
- [6] A. Contini, A. Canclini, F. Antonacci, M. Compagnoni, A. Sarti, and S. Tubaro, "Self-calibration of microphone arrays from measurement of times of arrival of acoustic signals," in *Int. Symp. Commun. Control and Signal Processing*, May 2012, pp. 1–6.
- [7] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino, "A closed form solution to the microphone position self-calibration problem," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Kyoto, Japan, Mar. 2012, pp. 2597–2600.
- [8] M. Crocco, A. Del Bue, and V. Murino, "A bilinear approach to the position self-calibration of multiple sensors," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 660–673, Feb. 2012.
- [9] I. Dokmanic, R. Parhizkar, J. Ranieri, and M. Vetterli, "Euclidean distance matrices: Essential theory, algorithms, and applications," *IEEE Signal Processing Mag.*, vol. 32, no. 6, pp. 12–30, Nov. 2015.
- [10] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. ACM*, vol. 24, no. 6, pp. 381–395, June 1981.
- [11] N. Gaubitch, W. Kleijn, and R. Heusdens, "Auto-localization in adhoc microphone arrays," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Vancouver, Canada, May 2013, pp. 106–110.
- [12] N. Gaubitch, W. Kleijn, and R. Heusdens, "Calibration of distributed sound acquisition systems using TOA measurements from a moving acoustic source," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Florence, Italy, May 2014, pp. 7455–7459.
- [13] M. H. Hennecke and G. A. Fink, "Towards acoustic self-localization of ad hoc smartphone arrays," in *Proc. Workshop Hands-Free Speech Commun. and Microphone Arrays*, Edinburgh, UK, May 2011, pp. 127–132.
- [14] M. H. Hennecke, T. Plötz, G. A. Fink, J. Schmalenstroer, and R. Haeb-Umbach, "A hierarchical approach to unsupervised shape calibration of microphone array networks," in *Proc. IEEE Workshop Statistical Signal Processing*, Cardiff, UK, Aug. 2009, pp. 257–260.
- [15] R. Heusdens and N. Gaubitch, "Time-delay estimation for TOA-based localization of multiple sensors," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Florence, Italy, May 2014, pp. 609–613.
- [16] F. Jacob, J. Schmalenstroer, and R. Haeb-Umbach, "DOA-based microphone array position self-calibration using circular statistics," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Vancouver, British Columbia, Canada, May 2013, pp. 116–120.
- [17] F. Jacob and R. Haeb-Umbach, "Coordinate mapping between an acoustic and visual sensor network in the shape domain for a joint self-calibrating speaker tracking," in *Proc. 11. ITG Fachtagung Sprachkommunikation*, Erlangen, Germany, Sept. 2014, pp. 1–4.
- [18] F. Jacob and R. Haeb-Umbach, (2015). Absolute geometry calibration of distributed microphone arrays in an audio-visual sensor network. *CoRR* [Online]. vol. abs/1504.03128. Available: <http://arxiv.org/abs/1504.03128>
- [19] F. Jacob, J. Schmalenstroer, and R. Haeb-Umbach, "Microphone array position self-calibration from reverberant speech input," in *Proc. Int. Workshop Acoustic Signal Enhancement*, Aachen, Germany, Sept. 2012, pp. 1–4.
- [20] A. Lauterbach, K. Ehrenfried, L. Koop, and S. Loose, "Procedure for the accurate phase calibration of a microphone array," presented at *Proc. 15th AIAA/CEAS Aero-acoustics Conf.*, Miami, FL, May 2009.
- [21] H. W. Löllmann, E. Yilmaz, M. Jeub, and P. Vary, "An improved algorithm for blind reverberation time estimation," in *Proc. Int. Workshop Acoustic Echo and Noise Control*, Tel Aviv, Israel, Sept. 2010, pp. 1–4.
- [22] S. Markovich-Golan, S. Gannot, and I. Cohen, "Distributed multiple constraints generalized sidelobe canceler for fully connected wireless acoustic sensor networks," *IEEE Trans. Acoust., Speech, Language Process.*, vol. 21, no. 2, pp. 343–356, Feb. 2013.
- [23] I. McCowan, D. Gatica-Perez, S. Bengio, G. Lathoud, M. Barnard, and D. Zhang, "Automatic analysis of multimodal group actions in meetings," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 305–317, Mar. 2005.
- [24] I. McCowan and M. Lincoln, "Microphone array shape calibration in diffuse noise fields," *IEEE Trans. Acoust., Speech, Language Process.*, vol. 16, no. 3, pp. 666–670, Mar. 2008.
- [25] N. Ono, H. Kohno, N. Ito, and S. Sagayama, "Blind alignment of asynchronously recorded signals for distributed microphone array," in *Proc. IEEE Workshop Appl. Signal Processing Audio, Acoustics*, New Paltz, NY, Oct. 2009, pp. 161–164.
- [26] M. Parviainen, P. Pertilä, and M. S. Hämäläinen, "Self-localization of wireless acoustic sensors in meeting rooms," in *Proc. Joint Workshop Hands-Free Speech Commun. and Microphone Arrays*, Nancy, France, May 2014, pp. 152–156.
- [27] P. Pertilä, M. S. Hamalainen, and M. Mieskolainen, "Passive temporal off-set estimation of multichannel recordings of an adhoc microphone array," *IEEE Trans. Acoust., Speech, Language Process.*, vol. 21, no. 11, pp. 2393–2402, Nov. 2013.
- [28] P. Pertilä, M. Mieskolainen, and M. S. Hämäläinen, "Passive self-localization of microphones using ambient sounds," in *Proc. Euro. Signal Processing Conf.*, Bucharest, Romania, Aug. 2012, pp. 1314–1318.
- [29] A. Plinge and G. A. Fink, "Online multi-speaker tracking using multiple microphone arrays informed by auditory scene analysis," in *Proc. Euro. Signal Processing Conf.*, Marrakesh, Morocco, Sept. 2013, pp. 1–5.
- [30] A. Plinge and G. A. Fink, "Geometry calibration of distributed microphone arrays exploiting audio-visual correspondences," in *Proc. Euro. Signal Processing Conf.*, Lisbon, Portugal, Sept. 2014, pp. 116–120.
- [31] A. Plinge and G. A. Fink, "Geometry calibration of multiple microphone arrays in highly reverberant environments," in *Proc. Int. Workshop Acoustic Signal Enhancement*, Juan les Pins, France, Sept. 2014, pp. 243–247.
- [32] A. Plinge and G. A. Fink, "Multi-speaker tracking using multiple distributed microphone arrays," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Florence, Italy, May 2014, pp. 614–618.
- [33] A. Plinge, R. Grzeszick, and G. A. Fink, "A Bag-of-Features approach to acoustic event detection," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Florence, Italy, May 2014, pp. 3704–3708.
- [34] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Las Vegas, NV, Mar. 2008, pp. 2445–2448.
- [35] V. C. Raykar and R. Duraiswami, "Automatic position calibration of multiple microphones," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Montreal, Quebec, Canada, May 2004, pp. iv–69–72.
- [36] V. C. Raykar, I. V. Kozintsev, and R. Lienhart, "Position calibration of microphones and loudspeakers in distributed computing platforms," *IEEE Trans. Speech and Audio Process.*, vol. 13, no. 1, pp. 70–83, Jan. 2005.
- [37] A. Redondi, M. Tagliasacchi, F. Antonacci, and A. Sarti, "Geometric calibration of distributed microphone arrays," in *Proc. IEEE Int. Workshop Multimedia Signal Processing*, Rio de Janeiro, Brazil, Oct. 2009, pp. 1–5.
- [38] J. Sachar, H. Silverman, and W. Patterson, "Microphone position and gain calibration for a large-aperture microphone array," *IEEE Trans. Speech and Audio Process.*, vol. 13, no. 1, pp. 42–52, Jan. 2005.
- [39] J. Schmalenstroer, F. Jacob, R. Haeb-Umbach, M. H. Hennecke, and G. A. Fink, "Unsupervised geometry calibration of acoustic sensor networks using source correspondences," in *Proc. Interspeech*, Florence, Italy, Aug. 2011, pp. 597–600.
- [40] J. Schmalenstroer, P. Jebrancik, and R. Haeb-Umbach, "A gossiping approach to sampling clock synchronization in wireless acoustic sensor networks," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Florence, Italy, May 2014, pp. 7575–7579.
- [41] M. J. Taghizadeh, R. Parhizkar, P. N. Garner, H. Bourlard, and A. Asaei, "Ad hoc microphone array calibration: Euclidean distance matrix completion algorithm and theoretical guarantees," *Signal Processing*, vol. 107, pp. 123–140, Feb. 2015.
- [42] M. Taseska and E. Habets, "Informed spatial filtering for sound extraction using distributed microphone arrays," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 22, no. 7, pp. 1195–1207, July 2014.
- [43] S. Thrun, "Affine structure from sound," in *Proc. Conf. Neural Informat. Process. Systems*, Vancouver, Canada, 2005, pp. 1353–1360.
- [44] S. D. Valente, F. Antonacci, M. Tagliasacchi, A. Sarti, and S. Tubaro, "Self-calibration of two microphone arrays from volumetric acoustic maps in non-reverberant rooms," in *Proc. Int. Symp. Commun., Control and Signal Processing*, Limassol, Cyprus, Mar. 2010.
- [45] S. D. Valente, M. Tagliasacchi, F. Antonacci, P. Bestagini, A. Sarti, S. Tubaro, and P. Milano, "Geometric calibration of distributed microphone arrays from acoustic source correspondences," in *Proc. IEEE Workshop Multimedia Signal Processing*, Saint Malo, France, Oct. 2010, pp. 13–18.
- [46] J. Velasco, M. J. Taghizadeh, A. Asaei, H. Bourlard, C. J. Martín-Arguedas, J. Macias-Guarasa, and D. Pizarro, "Novel GCC-PHAT model in diffuse sound field for microphone array pairwise distance based calibration," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Brisbane, Australia, Apr. 2015, pp. 2669–2673.
- [47] L. Wang, T. K. Hon, J. Reiss, and A. Cavallaro, "Self-localization of adhoc arrays using time difference of arrivals," *IEEE Trans. Signal Processing*, vol. 64, no. 4, pp. 1018–1033, Feb. 2016.

Hossein Nejati, Victor Pomponiu,
Thanh-Toan Do, Yiren Zhou,
Sahar Iravani, and Ngai-Man Cheung

Smartphone and Mobile Image Processing for Assisted Living

Health-monitoring apps powered by advanced mobile imaging algorithms

Smartphones are used by billions of people all over the world and are equipped with various types of sensors and increasingly powerful processors. The number of smartphone-savvy seniors is on the rise due to increased efforts to design elderly friendly smartphone systems and applications (apps). These smartphone capabilities, combined with advanced signal processing algorithms, are a

growing platform for different assisted-living solutions, ranging from human-computer interfaces for disabled users to health-monitoring and fitness-tracking devices. This combined software and hardware development, along with the culture of using portable and wearable devices, gives a practical, low-cost, and accessible solution for various assisted-living apps and, in particular, physiological monitoring for home or ambulatory settings [1]. However, many of these solutions are not yet mature enough to be released for public

Digital Object Identifier 10.1109/MSP.2016.2549996
Date of publication: 1 July 2016



use. Our focus here is health-monitoring apps based on mobile imaging to review the current state of the art and discuss the challenges.

Introduction

Mobile imaging is one of the main thrusts of smartphone apps for assisted living [2]. Today's smartphones are equipped with high-resolution image sensors. Typical smartphones can capture photos with a resolution of more than ten megapixels and significant image details. These camera-sensor advancements and the increased computational power enable smartphones to capture and analyze photos or videos for different apps, such as gestures and behavior recognition (for disabled users), object recognition for augmented reality purposes, contactless (and, therefore, noninvasive) preliminary self-diagnosis of diseases, self-monitoring of health conditions, and preliminary examinations. Increased research effort has recently been devoted to investigating novel health monitoring using smartphones and mobile imaging. Examples of these apps are heart and respiratory rate sensing [3], [4], eye examinations [5], wound assessment [6], and skin cancer detection [7]. All of these apps have been made possible through the ever-increasing number and type of sensors on the smartphones and the improvements in signal processing capability, and it is expected that new innovative smartphone-based apps will fill the remaining gap and make advanced health monitoring a publicly available component of assisted living.

Review scope

Focusing on signal capturing, processing, and analysis using smartphones as the platform, we investigate and survey how smartphones and signal processing are paving the way for a new unique class of assisted-living solutions: health-monitoring apps. We provide an overview of smartphone-based health-monitoring apps based on mobile imaging and discuss in detail the achievements and shortcomings in different categories of apps:

- the potential advantages of using smartphones and signal processing for health monitoring, with a review of several smartphone-based assisted-living solutions using various signal modalities (imaging, accelerometer signals, etc.)
- the unique advantages, design challenges, and considerations in using smartphones and mobile imaging for health monitoring
- the various image processing algorithms commonly used in different smartphone-based approaches
- three categories of health-monitoring apps using mobile imaging and their requirements and design considerations, processing algorithms, usability, and performance: 1) image-based heart rate estimation, 2) wound assessment and monitoring, and 3) preliminary skin cancer detection
- the open challenges and future research directions for using smartphones and mobile imaging in health monitoring.

Mobile imaging is one of the main thrusts of smartphone apps for assisted living.

Despite our focus on health-monitoring apps, our discussions are applicable to a wide range of assisted-living apps that either use similar underlying technologies and approaches or fall into one of the three categories of smartphone apps. Apps belonging to the same category have similar characteristics in terms of expected accuracy, computational complexity, energy consumption, and hardware requirements. We review image-based heart-rate-monitoring apps as a representative of low-risk assisted-living apps that are ready to be used by the public. Wound assessment and monitoring, representing assisted-living apps with midlevel risk, provide basic functionalities and may be used to assist less-experienced professionals or experts but are not ready for public use. Finally, a smartphone skin cancer diagnosis app represents high-risk assisted-living apps with the lowest level of maturity, not yet capable of providing assistance even to experts.

Signal processing for smartphone-based health monitoring: Advantages and design considerations

Much work has recently been dedicated to assisted-living apps on smartphones and signal processing, with increased attention on health-monitoring apps. Subjects include remote sensing, eHealth apps, sensor fusion, and smartphone-adapted diagnostic kits. In the past few years, smartphones have decreased in production cost and are now packed with different sensors, such as cameras, gyroscopes, accelerometers, and even pressure, altitude, and temperature sensors. In addition to the hardware, the culture of using smartphones suits health-monitoring apps. Smartphone owners carry their devices everywhere, making sure to keep the phones on 24/7. This culture therefore brings a unique, unprecedented opportunity for assisted-living apps and, in particular, health monitoring: access to several types of data related to the owner, with high temporal resolution and acceptable accuracy, throughout the lifetime of the phone.

The proper use of this unique opportunity, however, requires facing several challenges in both design and computational efficiency. Smartphones are made for everyday life, and users prefer a minimalistic design that conveys only important information, rather than being swamped with data. Examples are several one-button apps (e.g., the bSafe mobile app; www.getbsafe.com) with all of the processing hidden from the user and only necessary information and notifications revealed. Likewise, for an assisted-living smartphone app, the user should not be bombarded with data but monitored and assisted as needed. Because the assisted-living apps are almost always online and thus collecting more data, this design consideration should be noted with even more care. In addition, since many of these apps are targeted to the elderly population, usability issues should be carefully considered, for example, the reliability of image capture with unstable hand movement, the readability of the information to be conveyed even with poor eye sight, or the simplicity of choices and actions regarding the necessary steps to be taken. Finally, the collection of sensitive

information about the user calls for careful security-related designs in both on-device and on-server components of assisted-living mobile apps. This requirement is even more important for health-monitoring apps with imaging capabilities due to their access to a collection of private information about the user, including medical records and diagnostic results.

Assisted-living smartphone apps have important technical challenges in addition to the aforementioned design considerations: indirect measurements and computational, memory, and connectivity constraints. Health-monitoring apps have to be considered from two angles. On one hand, these apps are designed to perform some level of diagnosis based on the gathered information; therefore, incorrect or noisy information may lead to misdiagnosis, threatening the user's physical and psychological health. The challenge to acquire accurate data is more apparent in apps such as heart rate, walking distance, or respiration rate, where direct access to the targeted signal is impractical. A method has to be devised to estimate the signals from recordings of available sensors, such as cameras, global positioning systems, gyroscopes, and accelerometers. On the other hand, even with a mathematically sane method to estimate the targeted signal, providing a light-weight implementation of this method that complies with the computational power and memory limitations of the smartphone is one of the most difficult tasks in this field. For example, image processing algorithms play a significant role in modern-day health-monitoring apps, but their traditional implementation is for personal computers and cannot be directly used for smartphones. Smartphones are becoming increasingly powerful and are provided with more memory; nonetheless, the computational and memory costs of many commonly used techniques are still too high for smartphone capabilities.

In addition, to cover a wider range of users in different social groups, designers should avoid targeting merely the high-end, more powerful smartphones. To alleviate the computational and memory cost constraints, many mobile apps rely on on-server processing. However, assisted-living apps should not rely on the connectivity due to privacy and reliability issues. Data (e.g., images, locations, etc.) should be sent over a link for processing on the server, which is a vulnerable point for data leaks and attacks. There are also many possible scenarios in which a smartphone may fail to connect to a server when its services are vitally required. Therefore, designers should strike a balance between on-device (private and offline but slow and more basic) and on-server (fast and advanced analysis but requiring a good data link) processing to cover all scenarios with the best possible solutions.

Mobile image processing algorithms for assisted living

With a focus on assisted-living and, in particular, health-monitoring apps, we briefly review several commonly used algorithms for smartphone-based image processing: segmentation, feature selection, color analysis, and three-dimensional (3-D) structure reconstruction methods, with a brief overview of each method and its overall overhead (computational and

memory) on-device. Understanding the capabilities and constraints of different approaches will help an app designer to select the most appropriate method based on a target use and user population.

Image segmentation

Image segmentation uses the visual features of an isolated object or objects separated by boundaries to partition an image into several segments. In many assisted-living apps, this process is needed as one of the first steps to locate the region of interest (ROI). For instance, in skin cancer detection or wound assessment apps, the skin lesion or the wound boundary needs to be localized before other analyses can be performed [7].

Image segmentation methods can be divided into two categories: 1) boundary-detection-based approaches, which partition an image by discovering closed boundary contours, and 2) region-based approaches, which group together neighboring pixels with similar values and split groups of pixels with dissimilar values. The region-based methods usually represent an image as a graph, $G = (V, E, W)$, with the pixels as graph nodes V , and the pixels within distance r (graph radius) are connected by a graph edge in E . A weight value, $W(i, j)$, measures the similarity between pixels i and j . A higher $W(i, j)$ leads to greater similarity between pixels i and j . W can be computed using the location/illumination/texture information of pixels. The graph-based methods can be further divided into two subcategories. The first subcategory uses global information for segmenting. These methods are usually graph cut-based, e.g., normalized cut (Ncut) [8]. The second subcategory uses local information for segmenting, e.g., minimum spanning tree (MST)-based segmentation methods [9].

Although segmentation is a classic computer vision problem and has a wide range of state-of-the-art algorithms, image segmentation under simultaneous severe memory and computational constraints (e.g., performed on smartphones) has not been studied extensively. For instance, although Ncut can achieve good segmentation accuracy, it requires a considerable amount of computation and memory and therefore increases the overall energy consumption of the app. In contrast, MST-based segmentation methods have linear complexity, which does not dramatically increase the energy consumption of the app. However, because only local information is used to merge or split two segments, MST-based methods are usually sensitive to noise and perform poorly when applied to images captured by smartphones in uncontrolled environments. An example of the noise effect in the resulting segmentation is illustrated Figure 1.

The choice of the segmentation algorithm for each app depends on the segmentation difficulty in the app setting as well as the robustness of the entire system against errors in the segmentation step. For example, image-based heart rate estimation apps (see the "Image-Based Heart Rate Estimation" section) are significantly robust against errors in skin segmentation and can use a low-cost, color-based segmentation

approach. On the other hand, pixel-based apps, such as wound assessment (see the “Chronic Wound Assessment” section) or morphological image assessment, as in skin cancer detection algorithms (see the “Melanoma Detection Using Mobile Imaging” section), require a low error rate in the segmentation step because further processes are applied only to the segmented regions, or assessments heavily rely on the shape of the segmented region, as in the wound assessment and skin cancer detection apps, respectively. Therefore, these apps require an investment in more accurate and, in turn, more expensive segmentation algorithms.

Feature selection

Many of the advanced assisted-living apps, including most of the health-monitoring apps, include some level of machine-learning techniques. One of the considerations in designing an efficient machine-learning approach is feature selection, which can be performed offline to identify a small set of discriminative features to include in the smartphone-based apps. Feature selection is particularly important for smartphone-based analysis apps. Because of the computational and memory constraints of smartphones, it is imperative to identify a small and computationally inexpensive set of discriminative features to use for statistical model training and apps. Feature selection can reduce the time and memory requirements for feature extraction. In addition, feature-set reduction can decrease the model app time as well as the complexity of the classifier, resulting in lower overall memory and energy consumption of the app. In the case of health-monitoring apps, using a small set of features may also improve the classification performance because this reduces the chance of overfitting. For example, for skin cancer apps, a large number of features (regardless of computational and memory costs) may ideally encompass more information and help with correct diagnosis of a skin lesion. However, in reality, because the number of samples to train the required statistical model is usually low, the trained model will easily overfit unless a proper feature selection step is employed.

It is recognized that combinations of individually good features do not necessarily lead to a good classification performance. An exhaustive approach to find the best feature subset with, at most, k features from a set of M features ($k \leq M$) would require the examination of a significantly large number of feature subsets

$$\sum_{i=1}^k \binom{M}{i}.$$

In many cases, it is prohibitive to exhaustively search for the optimal feature subset.

A well-known practical feature selection procedure is the normalized mutual information feature selection (NMIFS) [10]. The NMIFS is a greedy algorithm that determines the optimal feature subset, starting with the feature that maximizes relevance with target class C . Given a set of selected features S_{m-1} , the next feature f_m is chosen such that it maximizes the relevance of f_m to target class C and minimizes

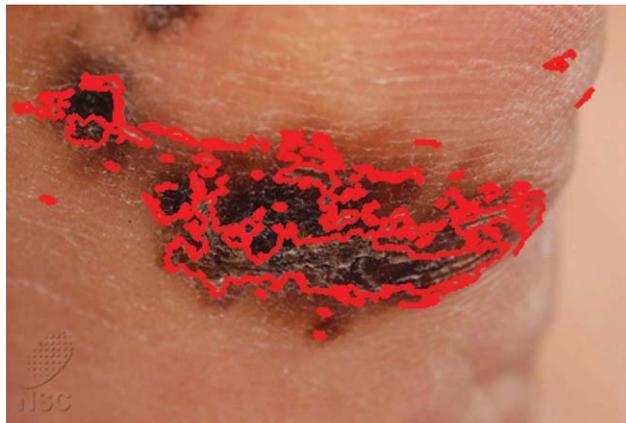


FIGURE 1. The skin lesion segmentation using an MST-based method [9]. Incorrect and separate segmented components may be due to image noise. (Photo courtesy of the National Skin Centre, Singapore.)

the redundancy between it and previous selected features in S_{m-1} . In other words, f_m is selected such that it maximizes G function

$$G(f_m) = I(C, f_m) - \frac{1}{|S_{m-1}|} \sum_{f_s \in S_{m-1}} NI(f_m, f_s), \quad (1)$$

where I is the mutual information (MI) function measuring the relevance between two variables and is defined as

$$I(X, Y) = \sum_y \sum_x p(x, y) \log \frac{p(x, y)}{p(x)p(y)},$$

and NI is the normalized MI function and is defined as

$$NI(X, Y) = \frac{I(X, Y)}{\min\{H(X), H(Y)\}},$$

where H is entropy function.

The additional constraints, such as required memory or computational cost, of each feature should be added to the optimization step to comply with the targeted mobile device constraints.

Overall, although feature selection is an expensive and time-consuming process, it can be performed offline on a powerful computer, and the final feature set is then used on the mobile device. Investing in a well-analyzed feature selection can significantly improve the speed, energy consumption, and even accuracy of the final app.

Color analysis

One of the challenges in image analysis is to find the appropriate color space for the problem being addressed. Given a proper color space, a segmentation or clustering algorithm may be used to select the ROI, followed by feature selection. Most of the systems developed for mobile image processing in the assisted-living field utilize color content

representation in red/green/blue (RGB) color images. The RGB information can be efficiently exploited for automatic image segmentation using, for example, K -means clustering family methods.

However, restricting the color information merely to RGB components is a simple abstraction that dismisses the information available within the color object. A color is fully defined by its complete wavelength response, whereas the RGB color space represents only three wavelengths. The perceived color also depends on the illumination condition, viewing angle, and sensor type. As a consequence, efficient color image processing requires an adequate color representation.

Different color spaces can be used to represent various color components, with different degrees of interdependency among them. Of the four classic color spaces, the hue/saturation/value (HSV), $YCbCr$, $L^*a^*b^*$, and RGB, the RGB color space has the most correlated components, whereas the $YCbCr$ color components are less correlated. This allows $YCbCr$ to extract uncorrelated components and favor the separation of the achromatic and chromatic parts. $L^*a^*b^*$ was originally designed to approximate human vision, with its L component (denoting the luminance) closely matching human perception of lightness. The color-opponent components are represented by a and b . One of the most important attributes of the $L^*a^*b^*$ model is device independence. This means that the colors are defined independently of their nature of creation or the device on which they are displayed. However, as the $L^*a^*b^*$ color space is much larger than that represented by RGB, it requires more data per pixel to reach the same precision as RGB. In contrast, the HSV color space is between RGB and $L^*a^*b^*$ in terms of balancing complexity and perceptual correctness. The HSV separates color into three components: two chromatic (hue and saturation) and one achromatic (value).

Based on the targeted app and the effect of color on components such as segmentation and automated decision making, a suitable color space can be selected. However, although most smartphone camera apps store the image in the RGB color space, the forward and backward transforms between RGB, HSV, and $L^*a^*b^*$ are not linear. One solution to the problem of conversion between different color spaces is to use a raw image format, but this solution adds extra overhead in terms of computation and storage.

In the end, similar to feature selection, color space selection is a time-consuming process that should be performed offline, and the decision to choose one space over the others should be based on the importance of color features and their robustness against image acquisition variations. For example, in the case of skin cancer detection, color features may not be as important as morphological features, and therefore, color normalization is not vitally important. In contrast, for an app such as wound assessment, the color

feature is one of the most important features for wound grading, so an investment in a more robust color representation may help the app perform better, although more slowly with a greater energy demand.

3-D computer vision and model

Three-dimensional computer vision uses the portability property of smartphones for better modeling. Because the user can easily move a smartphone, some information about the 3-D structure of the world is captured through the camera (video or image sequence) or other (usually add-on) sensors. In the context of health monitoring, 3-D structure and depth information is valuable in several apps. For example, wound depth is important to assess the condition of a wound, and the quantitative measurement of the elevation level of a skin mole is useful to detect an abnormal skin mole.

Three-dimensional computer vision uses the portability property of smartphones for better modeling.

Computer vision systems for 3-D scene reconstruction can be divided into active and passive approaches based on how the range (distance) information is captured. Active approaches, such as structured light, time of flight, and light detection and ranging, use measurements obtained from the

projection of a light beam onto a scene. The distance from the source to the scene at each pixel is measured by the travel of the light, producing the range (depth) images, in which the value of each pixel represents the calibrated distance between the camera and the captured scene. The benefits of using active systems are that they are less sensitive to illumination changes and provide range at every pixel location. However, the required equipment is normally available through add-ons, such as Structure Sensor (www.structure.io), and dramatically increase the cost and energy consumption of the app.

In contrast, passive approaches focus on the collection of images acquired by RGB cameras. One of the most prominent is the structure for motion (SfM), which uses the texture and color information of these images to find the corresponding 3-D point cloud and reconstructs the original scene. Although these methods do not require additional hardware, the main drawbacks are ambient light sensitivity and computational expense.

Unlike for large-scale structures such as buildings, passive range estimation techniques have not been studied extensively for small objects, including wound beds or skin moles. Therefore, we conducted an experiment to determine the accuracy of some representative passive range techniques for small objects. Figure 2 shows a reconstruction of the 3-D structure of a wound. We used a clay wound model comparable in size to a typical chronic wound (e.g., a foot ulcer) and incremental SfM [31]. Specifically, to obtain the 3-D reconstruction of the wound model using multiple photos captured from different viewpoints, the algorithm iteratively performs several tasks: feature computation, local correspondence matching, and fundamental matrix calculation. A bundle refinement with self-calibration was employed to compute the interior orientation with auxiliary parameters. We used scale-invariant feature transform features, which

are widely acceptable and relatively robust, to detect salient feature points in each image and then matched them in all other available images. This was followed by an app of random sample consensus to detect and remove quasi-degenerate cases. The remaining good matched points were then used to create the 3-D point cloud. The real model was 7 cm long, 4.5 cm wide, and 1 cm deep (at its deepest point), whereas the dimensions estimated from the 3-D reconstruction were 6.86, 4.51, and 1.24 cm, respectively (i.e., an approximately 0.24-cm error in the deepest depth). This passive 3-D reconstruction approach does not require additional hardware for smartphones, but does incur a significant computational and memory (and, therefore, energy) cost. Even if used for smartphone apps, this method lacks the execution speed that is required by many apps in this field.

As indicated by this experiment, passive techniques may not provide accurate depth information (e.g., for wound depth), although passive models could be used for two-dimensional (2-D) object dimension estimation (e.g., 2-D dimensions for wound or skin lesions). This error in estimating the depth is usually due to fewer point correspondences with a limited view of these parts. In contrast, the range values obtained in active approaches are more accurate due to precise calibration of the equipment used, but the 2-D spatial resolution achieved by active approaches is much lower than that obtained by passive approaches.

Table 1 provides a brief overview of 3-D depth map reconstruction methods in terms of methodology, accuracy, resolution, range, sensitivity to scene illumination and surface features, computational cost, usability and portability, and the hardware/software costs. Different approaches for 3-D reconstruction are associated with different overheads and may require different add-ons. For example, compared with passive approaches, active approaches are robust against illumination variations but are less portable and, therefore, suitable for different problem settings.

As for the currently available mobile hardware, the cost of using 3-D modeling is still too high and can be justified only if the app strongly relies on 3-D modeling. For example, in apps such as wound assessment that require accurate depth estimation and can tolerate high energy consumption, low-level passive or mid/high-resolution active 3-D modeling seems suitable. In addition, as we discuss in the next section, based on recent advancements in mobile hardware technology, such as the addition of visual computing units (VPUs), running passive 3-D modeling approaches on mobile devices may be possible.

Hardware requirements

When deciding on the approaches to be incorporated into a smartphone app, two main characteristics of the target problem should be considered: 1) the data modality required by the app and the benefits and consequences of each family of approaches (e.g., 2-D or 3-D modeling) and specific methods (e.g., active or passive 3-D reconstruction) and 2) the hardware requirements of any chosen approach and the resulting overhead for the system. Many commonly used methods, such as

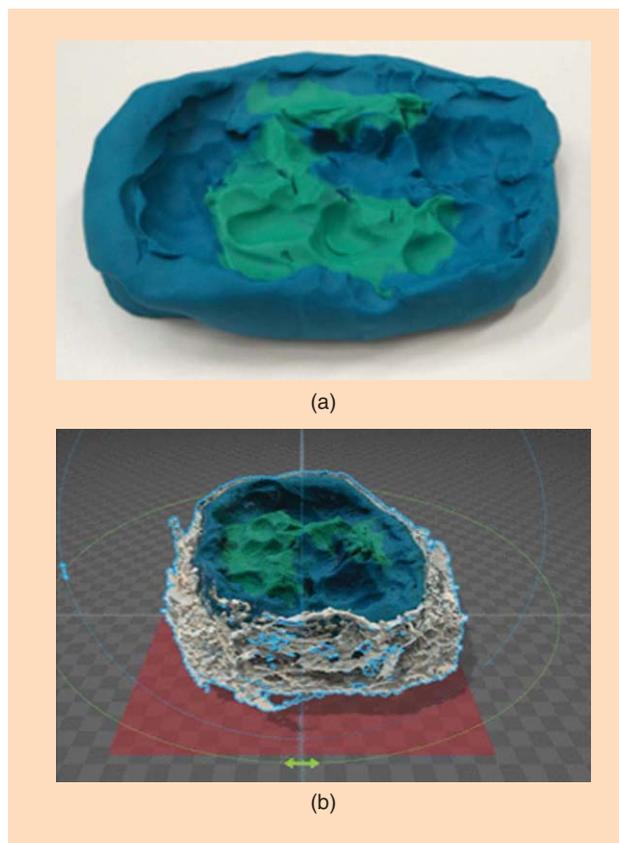


FIGURE 2. (a) A clay model of the wound and (b) our 3-D reconstruction of it (model dimensions: 7 cm long, 4.5 cm wide, and 1 cm deep at the deepest point.)

Table 1. A comparison of 3-D depth map reconstruction approaches.

	SfM	Stereo	Time of Flight	Structured Light
Vision system	Passive	Passive	Active	Active
Reconstruction accuracy	++ ^(a)	+++	+	++
Distance (spatial) resolution	+++	++	+	+++
Range	++	+	+++	++
Lighting sensitivity	+++	+++	-	-
Material sensitivity	+	++	-	+++
CPU/GPU utilization	++	+++	+	++
Usability	+++	++	++	+
Portability	+++	++	+	+
Cost	+	++	+++	++

^(a) -: none; +: low; ++: medium; +++: high.

segmentation, or simple classification techniques work on the common components of smartphone platforms without requiring special configurations [e.g., onboard graphics processing units (GPUs)] or additional hardware add-ons (e.g., 3-D scanner add-ons). Therefore, if used for a smartphone app, they would generally work similarly across generic smartphones.

However, when an approach requires specific hardware capabilities, the performance of the system can greatly vary based on the differences in the related hardware.

The hardware difference issue may not be widespread with regard to specialized add-ons (due the limited number of options), but this performance difference problem is also present even in a hardware component as common as a camera. In [11] and [12], camera phone capabilities were compared in depth with consumer digital cameras with respect to camera sensors, lens distortion, reliability, stability, and robustness for a particular image processing task. In [12], 16 mobile phone cameras and 50 digital cameras were evaluated for close-range 3-D reconstruction. The experimental results show that, on average, the capabilities of phone cameras with at least eight-megapixel resolution for 3-D reconstruction are similar to those of digital cameras. Other smartphone camera aspects of 2-D imaging have also shown to be similar to handheld digital cameras in other app-specific cases, such as analyzing skin images, as discussed in the section “Chronic Wound Assessment.”

The phone camera displays a different lens distortion effect, with more distortions to the edges of the image than to the center. This phenomenon can be very difficult to handle for the mobile cameras due to their inherent design of small optics and large fields of view. The environmental conditions also have a strong impact on the quality of the captured images. In particular, the illumination conditions affect the exposure, color, and focus of the acquired image. All of the mentioned hardware characteristics of the mobile phone can affect the captured image and, in turn, the processes that use the image signal as their input (e.g., many medical apps). For instance, in an uncontrolled environment, the registered color for the skin depends on the spectral power distribution of the light source and the spectral reflectance of the tissues. This change in color has a significant effect on many preprocessing steps (e.g., color-based segmentation) and advanced steps (e.g., color-based feature extraction and classification). As we discuss in the next section, due to the complexity and computational costs of illumination and color normalization, many apps, such as wound assessment and skin cancer diagnosis, suffer from variations in these factors.

There are also some differences in how the signal captured by the camera sensor (raw image) is interpreted into typical color formats such as JPEG. These difference are similar to color interpretation differences in normal digital cameras. Therefore, the similarity of smartphone cameras and consumer digital cameras reported [11] comes with the caveat of high intraclass variability in the generated 3-D reconstructions. To cope with these intraclass variability issues, standard protocols, such as the IEEE P1858-2012 camera phone image quality (CPIQ) standard [13], have been proposed to define the CPIQ metrics. The IEEE CPIQ standard is an ongoing effort to standardize image quality across image acquisition devices, with version 1 published in 2016 (<http://grouper.ieee.org/groups/1858/>). This standard includes seven quality metrics:

- spatial frequency response
- lateral chromatic displacement
- chroma level
- color uniformity
- local geometric distortion
- visual noise
- texture blur.

For assisted-living apps, mobile devices should have competent quality for these metrics. For example, tissue color is an important feature for a wound assessment app, and the chroma level (colorfulness) and color uniformity of the devices should be competent. Likewise, border irregularity is an important feature for a skin cancer prescreening app, and the spatial frequency response (resolution, sharpness) of the devices should be superior. Detailed analysis of these quality metrics should be carried out for assisted-living apps upon the official release of the IEEE CPIQ standard.

Finally, it is important to note that, due to the continuous advancements in mobile hardware technologies, the computational power gap between mobile devices and personal computers is rapidly closing. This results, in turn, in the applicability of powerful, computationally expensive algorithms on mobile devices without compromising other device services. Examples of these advancements are the prospect of the addition of VPUs to Android phones under a partnership between Google and Movidius [32]. The use of VPU chips will enable smartphones to locally run advanced image processing algorithms, such as state-of-the-art deep learning, 3-D depth, and spatial computing, with significantly lower energy requirements than currently possible. It seems feasible that many of the algorithms discussed can be run on mobile devices in the near future.

Examples of designs and challenges

In this section, we present three categories of smartphone apps for health monitoring in assisted-living scenarios: heart rate monitoring, wound assessment, and skin cancer diagnosis. These apps represent three important categories in this field in terms of input and expected accuracy in the output. Heart rate is a daily monitoring app for a wide range of purposes (e.g., sports tracking or health-care-related monitoring) that require real-time analysis in an uncontrolled environment. Wound assessment using smartphones is a monitoring app for regular and time-based usage. Although it requires advanced processing techniques, it uses a relatively controlled image acquisition procedure, and inaccurate results are acceptable to some extent. Finally, the smartphone skin cancer diagnosis app is increasingly popular and requires advanced image processing and pattern-recognition techniques. In addition, the results produced by this app should be particularly accurate due the possibility of inflicting severe damage to the user's physical and psychological conditions. It is worth mentioning that this automatic skin cancer diagnosis is still an unsolved problem even without smartphone platform restrictions. Therefore, developing a desirable skin cancer detection app is the most challenging task in these three categories. These three apps are

Table 2. A comparison of smartphone-based heart-rate-monitoring approaches.

	Motion-Based Methods		Video-Based Methods	
	Single Sensor [15]	Multisensor [14]	Fingertip Video [16]	Video Magnification [17]
Accuracy	+(^a)	+++	+++	+++
Robustness	+	+++	++	+
Real-time	+	+	+	-
Energy consumption	+	++	++	+++
Practicality	+	+	++	+

^(a) -: none; +: low; ++: medium; +++: high.

representative of their app category in health monitoring and also broader assisted living.

Image-based heart rate estimation

Heart rate is traditionally measured based on changes in the pressure, or electrical activity of the heart [electrocardiograph], sensed at the skin level. Heart rate can vary according to the body’s physical requirements and conditions, including absorbing oxygen, excreting carbon dioxide, exercising, sleeping, ingesting, anxiety, stress, illness, and drugs. Different approaches using various sensor modalities have been used for smartphone-based heart rate estimation (summarized in Table 2). Based on the nature of the heartbeat, some methods [14], [15] have proposed the use of sensors such as accelerometers and gyroscopes to estimate the heart rate and can decompose the input signal into heart rate and respiratory signals. Additional statistical analyses have also led to differentiation between types of physical load (e.g., walking, running, etc.) and accordingly regulate the heart rate estimation module. The profiling approach has been extended to differentiation between groups of people (e.g., based on their level of athleticism). Despite the benefit of multimodal sensory input and possibility of (ideally) estimation of different physiological

signals, the input signal to these approaches can be collected only by contact sensors, which limits their applicability and ease of use.

To remotely estimate the heart rate, Poh et al. [3] proposed an algorithm that analyzes subtle color changes in a user’s face image, captured either via a webcam or a smartphone camera (from the reflection image in a mirror). In this method, independent component analysis (ICA) is applied to the normalized time series from the RGB color components of the captured video. Analyzing the power spectra of the resulting three independent recovered source signals accurately estimates the heart rate (Figure 3). The accuracy of experimental results with both stable and moving bodies has been reported to be significantly higher than that achieved using thermal imaging of the major superficial vessels.

The heart rate estimation research has been extended for use on smartphones in several other studies, including [4], [16], and [17]. Kwon et al. [4] proposed a phone-based heart rate measurement method that analyzes color changes in a video of a face captured by a smartphone camera. Experiments showed a 1.04% error rate in a small group (ten people) of subjects. This method is, however, sensitive to motion artifacts and ambient light interference. The motion artifact is taken into

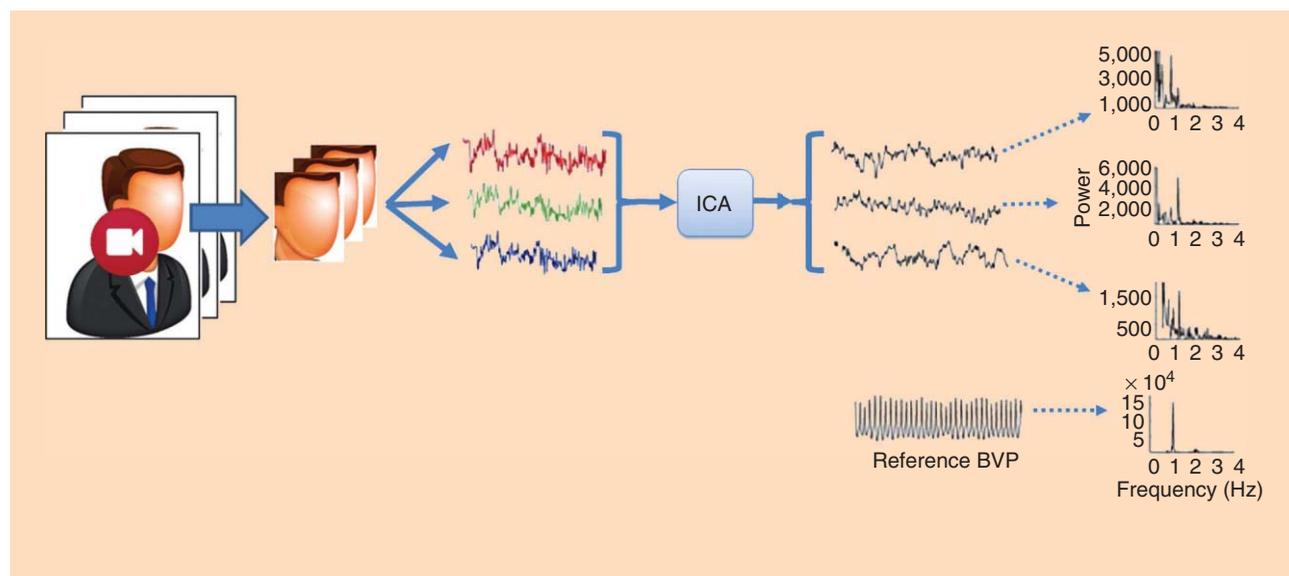


FIGURE 3. Estimating heart rate based on the ICA of the color channels of a face video captured by a smartphone camera [3]. BVP: blood volume pulse.

account by Sun et al. [18], who introduced planar motion compensation and blind source separation to stabilize the input video sequence before use by the signal detection module. Pal et al. [16] introduced a more robust method that uses a phone camera to record and further analyze small color changes on a fingertip. Instead of manual skipping of the first few seconds (which are usually noisy due to finger movement) used in previous works, in [16], the authors introduced a finite-state machine (FSM) that employs multiple-window short-time fast Fourier transform to detect variations in the color of a fingertip placed on a phone camera with the flash on. The authors also proposed a video signal rejection method based on color filtering and heuristic thresholds of the mean red value μ_R to avoid feeding noisy videos to the heart rate detection module. For the frame rejection FSM, Pal et al. [16] used a Mealy machine, which determines the output based on both the machine state and the input, while the next state is also governed by the current state and the input value. Based on this FSM, the algorithm tries to capture α consecutive frames with mean red value $\mu_R > \sigma$ or otherwise reset the frame-capturing routine. When enough frames pass this criteria using 512-point fast Fourier analysis, the algorithm tries to find peak frequencies of the red value in the frames. If the peak frequency is out of the normal human heartbeat rate range or the variance of retrieved frequencies is too wide, the set of frames can be rejected, and the algorithm restarts from the frame-capturing routine. Finally, having passed all of these criteria, the peak frequency of the set of frames will be the output of the algorithm, as the estimated heartbeat rate. Based on the preprocessed signal, this method reveals several physiological variables, including breathing rate, cardiac R-R intervals, and blood oxygen saturation. By comparing their estimated readings with standard methods for making such measurements (respiration belts, electrocardiograms, and pulse oximeters, respectively), the authors showed the practicality of this app as well as the possibility of using the same set of techniques to detect atrial fibrillation or blood loss.

This method accurately estimates the heart rate as well as detects the beat variation, given a good-quality video sequence. Several mobile apps have been developed based on this work for the Android, iOS, and Windows Phone, including Instant Heart Rate (www.azumio.com/s/instantheartrate) and Heart Beat Rate (www.heartbeatrate.com), and have even been developed as a built-in feature for the Samsung Galaxy S5 and S6 series. Due to relying only on relatively simple algorithms, this approach has a relatively low computational cost and energy consumption. However, the drawback of the use of extensive preprocessing is the lack of any input when any of the aforementioned noise occurs. Even when the preprocessing step accepts the frames, the readings can vary based on how the finger is placed on the camera sensor. This drawback may reduce the practicality of the approach and the accuracy of the results when used without assistance or with an unstable hand (e.g., during exercise or when used by elderly people).

Xu et al. [17] published one of the most recent works on heart rate monitoring using a phone camera based on the amplification of color changes in a face video caused by blood

being pumped in and out of the face. Because these subtle color changes are at the same frequency as the heartbeat, the amplified changes can be used for heartbeat rate estimation [17]. This magnification is achieved by temporal filtering, followed by spatial filtering, which are both tuned to specifications of the human heartbeat (e.g., 0.4–4 Hz and the importance of the color red over green and blue). The main idea is to use Taylor series expansion to estimate changes in the signal (here, the color of a given face area over time). If the value at pixel x at time t is given by $I(x, t) = f(x)$ and at time $t + 1$ by $I(x, t + 1) = f(x + \delta)$, then this change can be approximated by Taylor series expansion as

$$I(x, t + 1) \approx f(x) + \delta(t) \frac{\partial f(x)}{\partial x}. \quad (2)$$

Let $B(x, t)$ be the temporal filter tuned for the targeted change frequency, here the human range of a heartbeat rate,

$$B(x, t) = \delta(t) \frac{\partial f(x)}{\partial x}. \quad (3)$$

Therefore, we can reach an amplified signal $\tilde{I}(x, t)$ by scaling this change by α and add it back to the original signal $I(x, t)$

$$\tilde{I}(x, t) = I(x, t) + \alpha B(x, t). \quad (4)$$

Combining (2)–(4), we have

$$\tilde{I}(x, t) = f(x) + (1 + \alpha) \delta(t) \frac{\partial f(x)}{\partial x}. \quad (5)$$

As illustrated in Figure 4, using this approach, we amplified the changes in the red channel of a video sequence captured by a normal smartphone held in a stable hand. Using the bandwidth of the human heartbeat range, we prevented the amplification of the signal noise. A simple frequency analysis of this amplified change over time proves the method to be as accurate as clinical heart-rate-monitoring instruments [17].

This method provides a contactless method of heart rate estimation, allowing third-party use (e.g., caretakers or nurses) and applicability to a wider range of users (e.g., infants and seniors). However, it is still not robust against motion artifacts and requires the camera and the subject to be relatively stable. In addition, due to the higher-complexity order of the algorithms used, the computational cost, memory requirements, and energy consumption of this app are still too high for midlevel smartphones.

Table 2 provides a summary of smartphone-based heart-rate-monitoring apps and an overview based on their methodology, achieved accuracy, robustness to noise, execution time, energy consumption, and overall practicality. Single-sensor approaches have the benefit of low-power dissemination, a characteristic that is vital for many continuously monitoring apps. However, because the single-sensor approaches do not provide accurate results, they are suitable for noncritical apps, such as general fitness tracking. More accurate solutions come with either higher power consumption (e.g., for multisensor or fingertip video approaches) or a requirement for a controlled environment (e.g., for video magnification

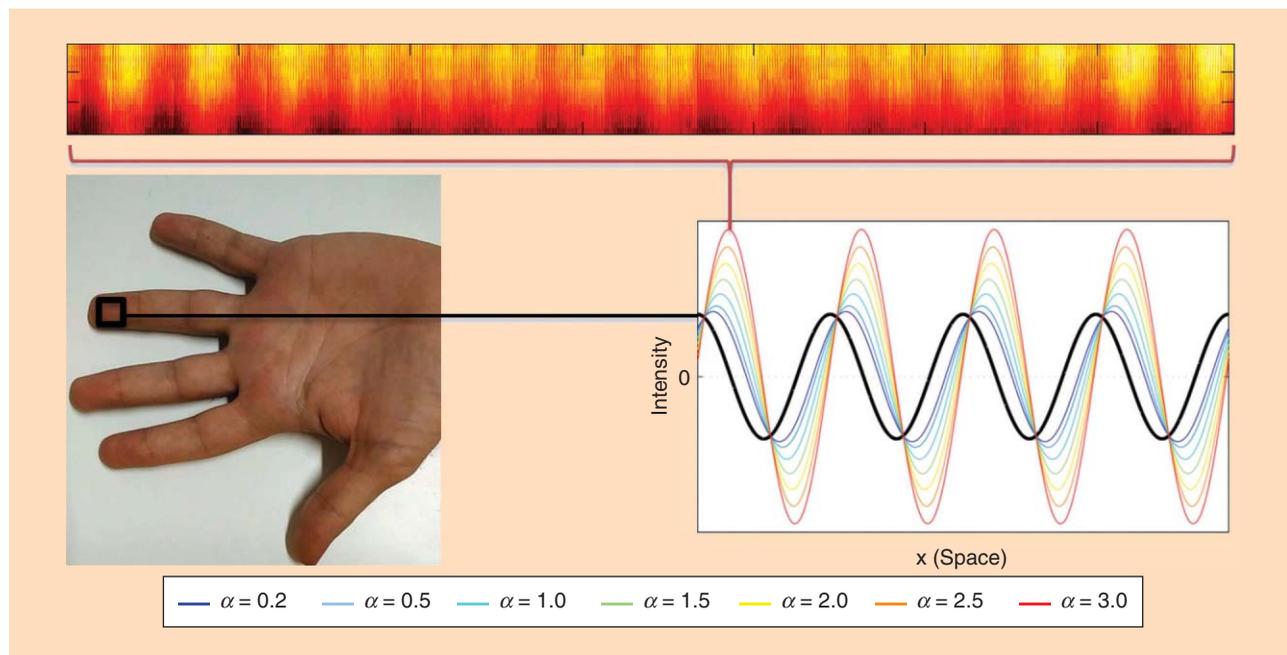


FIGURE 4. An example of the signal change magnification from a smartphone-captured video of a stable hand to obtain the heartbeat rate (based on [17]).

approach). A smartphone app therefore can use a wide range of solutions based on goals and problem constraints. For example, a quarantined patient can be remotely and accurately monitored by the video magnification approach to avoid contact because power consumption is unlikely to be a pressing issue.

Chronic wound assessment

Chronic wounds, such as pressure ulcers and diabetic wounds, are a drain on health-care resources and, as such, continue to challenge health-care providers to define and create more effective intervention strategies. Wound assessment is the foundation for maintaining and evaluating a therapeutic plan of care. Without adequate baseline wound assessment and valid interpretation of the assessment data, the plan of care may be inappropriate or ineffective [19].

To reliably assess the wound type, grade (severity), and healing process, accurate and objective measurements, such as the area, depth, and tissue composition, are needed. Primary measurements, including width, length, and depth, which are critical for assessing wound severity, can be obtained through noninvasive and visual methods. In contrast, more complex measurements, such as tunneling and undermining, are determined with invasive methods (e.g., using a cotton swab). In addition to wound dimensions and composition, trained specialists use further information, including inflammation, temperature, exudate levels, and periwound condition (i.e., the surrounding skin). Based on this information and years of experience, the specialist then estimates the type and grade of the wound and proceeds with an appropriate treatment. Therefore, it is logical to take all of these information modalities into account for an automatic wound assessment system. The framework that seemingly encompasses these considerations is presented in Figure 5.

The challenge of including all of these signal modalities is evident with regard to a smartphone app. This app should comply with the clinical evaluation of the wound (i.e., on what to measure and estimate) and with clinically acceptable results to integrate into the rest of the health-care system. To be considered an effective tool in the hands of caregivers, the app should provide a faster and safer method than traditional methods to assess and track wounds. Moreover, privacy issues should be carefully examined, in addition to other on-device processing considerations. A well-designed and accurate smartphone-based solution for wound assessment would dramatically reduce wound-care costs and difficulties both for patients and caregivers. Due to the capabilities of the smartphone platform, the presumed smartphone app can also be easily integrated into the health-care ecosystem, providing functionalities such as telemedicine and location-based emergency support systems.

Handheld solutions for wound assessment and monitoring can be divided into two categories, with the apps in the first category only estimating the wound size (i.e., the area and/or volume) and the apps in the second category performing tissue classification to estimate the wound condition. We discuss image-based mobile solutions for noninvasive wound assessment and review scientific literature and the apps markets for Android and iOS platforms. Table 3 compares the families of solutions, each with an example.

Wound size estimation methods

Nonautomatic methods for wound size estimation are time-consuming and subjective and suffer from high variability. Therefore, semiautomatic and digital planimetry techniques have become popular and are based on the digital images of a wound. Proposed solutions, including Wound Tracker

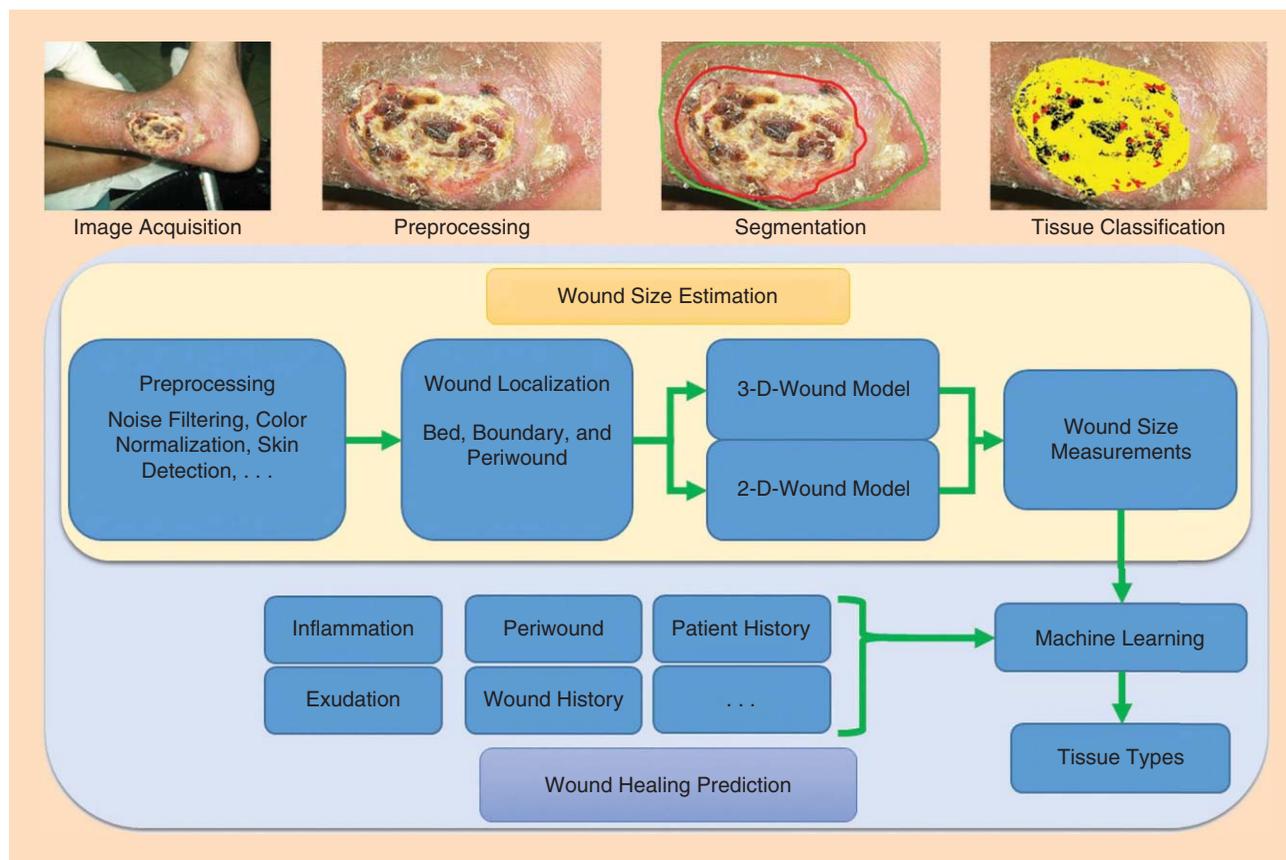


FIGURE 5. The general processing framework for the quantitative assessment and monitoring of wound healing using a smartphone. (Photos used courtesy of Wikipedia.)

Table 3. A comparison of the wound assessment methods.

	Wound Localization	Measurement		Tissue Analysis	Processing	Hardware	Usability	Portability
		2-D	3-D					
Wound Tracker	Manual	Manual	— ^(a)	M	Local	Markers	+	++
WoundMAP	Manual	Manual	—	Automatic	Local	Ruler	+	+
MOWA	Manual	Manual	—	Automatic	Server	Markers		+
Wound Analyzer	—	—	—	Semiautomatic	Local	Markers	+	+
Wound Analysis	Automatic	Manual	—	—	Local	Markers	+	+++
TissueAnalytics	Manual	Manual	—	Automatic	Server	Markers	++	++
Hettiarachchi et al. [20]	Semiautomatic	Automatic	—	—	Local	Markers	+	+++
WoundManager	Semiautomatic	Automatic	Semiautomatic	Automatic	Server	Special Equipment	++	++
Wang et al. [21]	Automatic	—	—	Automatic	Local	Special Equipment	+	+

^(a) —: none; +: low; ++: medium; +++: high.

(www.woundtrackerpro.com), Wound Analysis (www.woundanalysis.wordpress.com), WoundMAP (www.itunes.apple.com/us/app/woundmap/id673582282?mt=8), and MOWA (www.healthpath.it/imowa.html), use a smartphone (or a handheld device) to manually sketch the contour of the wound on the touchscreen of the device to calculate the area.

For this purpose, a ruler, markers, or other accurate scaling equipment is placed in the vicinity of the wound before acquiring the photo. These methods perform minimal computations but require heavy human input.

A more automatic method to estimate 2-D wound dimensions was proposed by Hettiarachchi et al. [20] based on a

dynamic active contour model. The 2-D wound measurements are computed automatically after the calibration of the camera phone. But the method is sensitive to the viewing angle and distance variations, which highly affect the usability of the proposed scheme.

To measure the area and the volume of a wound, 3-D models of the wound bed are needed. Tracking these 3-D measurements over time may provide useful diagnostic information for wound grading and healing process estimation. Recently, volumetric methods were devised to obtain more accurate measurements. WoundManager (www.woundmanager.com), for example, computes the 3-D structure of a wound using a 3-D camera that can be attached to a mobile device. The user first selects the ROI containing the wound. Second, the information from the range (3-D) image is used to automatically complete the wound segmentation and obtain the wound dimensions. The human input helps the system by removing unwanted artifacts from the image and is essential for the next processing steps. Although the measurements obtain from 3-D models can result in more accurate and reliable analysis, as discussed in the section “3-D Computer Vision and Model,” these methods are not yet optimized for smartphone-based apps.

Wound tissue classification methods

In addition to regular tracking of the wound size, tissue classification is a vital aspect that can guide clinicians in recommending the most suitable treatment. The most common types of wound tissue are necrotic, sloughy, healthy granulating, unhealthy granulating, hypergranulating, and epithelializing [22]. In cases in which the granulating process is affected by a disturbing factor (e.g., infection), the tissue is unhealthy and granulated with a darker red color and is sometimes covered with a jelly-like substance. If the unhealthy granulation continues, the tissue will start to overgrow and severely affect the periwound. This condition is referred to as hypergranulation. Treatments for unhealthy and hypergranulating tissue should be determined based on the cause, which can be reaction to external materials (such as tubes or metallic parts) or bacterial infection.

Early apps such as WoundMAP, MOWA, and Wound Analyzer (www.appcrawlr.com/ios/pressure-wound-analyzer) proposed a simple correlation rule between the evolution of the healing and the dominant color within the wound bed. Improved approaches such as AWAMS (Savant Imaging, New York) calculate the percentages of tissue types and the area of the wound after the user manually separates the tissue types. Other than requiring extensive manual input, these approaches are sensitive to the camera viewpoint of the wound.

More advanced algorithms employ pattern-recognition techniques for automated tissue classification. In several studies, wound tissue image samples were manually labeled to train a statistical model for tissue clustering and classification. Recently, unsupervised tissue classification in different color spaces was also investigated [22].

In addition to color-based features, several other image-based features, such as Gabor coefficients, local binary patterns (LBPs), co-occurrence matrices for texture contrast, and

anisotropy, have been employed in wound tissue classification. But the color-based features are still the most discriminating and are usually sensitive to uncontrolled lighting conditions, a common condition for smartphone-based imaging. Therefore, a common practice is to include a colored checker pattern in the field of view of the camera and then perform color normalization as a preprocessing step. In addition to the colored checker pattern, a more complicated setup has been proposed by Wang et al. [21] to reach color constancy. In this approach, the ambient light is blocked by a cover, and wound images are obtained using only a light with known parameters. Although this work has shown accurate classification of tissue types, the required setup and auxiliary hardware dramatically reduce the practicality of this approach.

Regardless of the reported performance of the wound tissue classification methods, they cannot be used for clinical assessment because the proposed solutions found in the literature and smartphone app markets consider only three tissue classes: black (for necrotic), yellow (for slough), and red (for granulating). In other words, these apps entirely ignore epithelializing tissue and combine all three types of granulating tissues into one class. Figure 6 compares healthy and unhealthy granulation, the two classes that are regarded as one in the existing



FIGURE 6. (a) Unhealthy and (b) healthy granulating tissues. (Photos courtesy of Bright Vision Hospital, Singapore.)

works. This failure to comply with the clinical methodologies of wound assessment have rendered these apps incompetent and, more important, misleading.

The shortcomings of the existing methods using color and texture descriptors are illustrated by comparing their performance in a three-class scenario (black, yellow, and red) and in a five-class scenario (necrosis, slough, healthy granulation, unhealthy granulation, and hypergranulation). Using a database of 1,000 wound tissue images, we simulated the main approaches in the literature. We used the RGB and HSV histograms as color descriptors and the LBP as a texture descriptor. Color and texture features were fed to the classifiers both separately and in combination. We used a support vector machine (SVM), K -nearest neighbor (KNN), and K -means as representative classifiers. The results of this experiment are summarized in Table 4 and indicate that, although the features revealed a high discriminative power (highest with the SVM) in the three-class scenario, they failed to reach an acceptable level of performance when tested in the five-class scenario and thus cannot be used for clinical purposes.

Melanoma detection using mobile imaging

Malignant melanoma is the most aggressive form of skin cancer, forming in the pigment cells of the epidermis, and is responsible for the majority of skin-cancer-related deaths. However, most cases are curable if they are detected at an early stage. This fact is the main motivation behind several attempts to develop a smartphone-based skin cancer detection app that is easy to use, frequently accessible, and capable of providing accurate and reliable diagnostics. A skin cancer detection app may seem to have challenges similar to a wound assessment app. This is true, to some extent, with regard to image acquisition, normalization in an uncontrolled environment, and performing all processing with smartphone capabilities, privacy and security issues, and scalability of the provided solution. However, although use by the general public is the ultimate goal of this specific smartphone app for inexpensive, easy, and frequent assessment of skin moles, there is a critical distinction between this

app and heart-rate-monitoring or wound assessment apps: a wrong skin cancer diagnostic result can have a catastrophic effect on a person's physical and psychological conditions. The challenge of developing an accurate skin cancer detection app is more evident when considering the target user group—the untrained public. Nonetheless, several works have been proposed to improve the state of the art, starting from more constrained scenarios.

A general categorization of melanoma diagnosis methods can be based on the level of required human input. On one end of the spectrum are manual methods, such as Lubax (www.lubax.com), that minimally use smartphone platforms (e.g., for transferring images) and require the visual inspection of an experienced dermatologist. On the other end are fully automatic methods [7] that perform the assessment without human intervention (besides steps such as image acquisition). Finally, hybrid approaches [23] are fall between these two extreme methods and usually provide some automatic assessment of the skin lesion based on visual and context information (e.g., skin type, age, gender, family history, etc.) to help dermatologists make their decisions. An illustration of the generic methodology that should be employed for smartphone-based skin cancer diagnosis apps is shown in Figure 7. The framework indicates the assessment of both local and contextual information, in addition to typical preprocessing steps. This framework suggest a continuous monitoring and recording of the lesion status, which is a clinically emphasized approach. Apps that are proposed or are currently available in the market may have some (but not all) components of this generic methodology based on their goal of what problem to solve in skin cancer detection.

Fully automatic and hybrid melanoma diagnosis systems in the literature usually have three main phases. The first phase is image acquisition, which can be performed with different devices, such as dermatoscopes, spectroscopes, standard digital cameras, and smartphone cameras. The images acquired by these devices exhibit different qualities that can significantly change the outcome of the analysis process. The second phase is ROI selection, which usually involves skin detection, removal of unrelated image sections (e.g., ruler,

Table 4. Our results for wound tissue classification based on three-class (black, yellow, and red) and five-class (necrosis, slough, healthy granulation, unhealthy granulation, and hypergranulation) scenarios using the main approaches in the literature.

Feature	Number of Classes	SVM			K-Means			KNN		
		Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
RGB H	Three	0.8500	0.8513	0.850	0.3700	0.38093	0.3700	0.812	0.816	0.816
	Five	0.7710	0.7690 ^(a)	0.769	0.3610	0.2546	0.3610	0.731	0.742	0.731
HSV H	Three	0.7600	0.7675	0.760	0.5983	0.67908	0.5983	0.7683	0.7758	0.7683
	Five	0.675	0.6898	0.675	0.273	0.2758	0.273	0.648	0.664	0.648
LBP	Three	0.785	0.7842	0.785	0.3033	0.3547	0.303	0.870	0.8702	0.870
	Five	0.675	0.6904	0.675	0.2650	0.2575	0.265	0.645	0.6630	0.645
RGB+LBP	Three	0.8766	0.8694	0.877	0.3750	0.43756	0.3750	0.8250	0.8179	0.8265
	Five	0.7650	0.7663	0.765	0.2500	0.2557	0.2500	0.7340	0.7441	0.7340

^(a) The best accuracy for the five-class scenario is only 0.7690. Therefore, existing works fail to accurately differentiate wound tissues that have similar appearances but require different treatments, e.g., healthy and unhealthy granulating tissues.

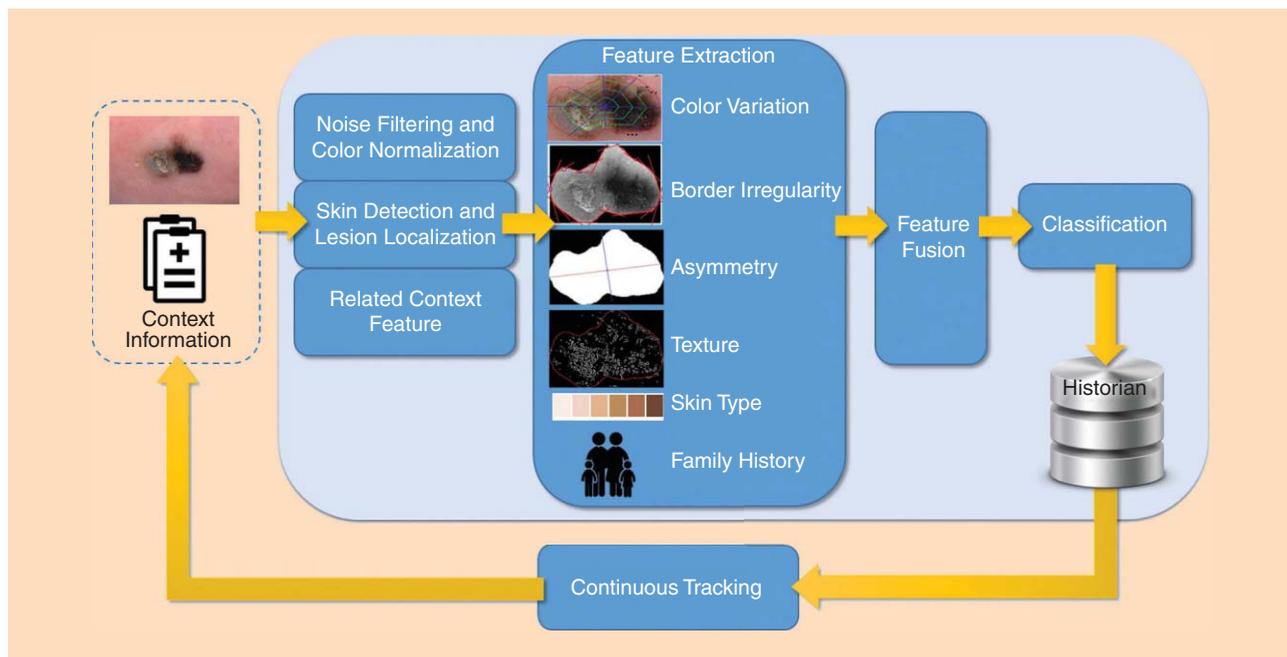


FIGURE 7. The generic methodology used in holistic and continuous monitoring for smartphone-based skin cancer diagnosis. (Skin lesion photos courtesy of the National Skin Centre, Singapore.)

watch, hair, scar), and mole border localization. Finally, the third phase is mole classification, which includes feature extraction and use of a statistical model to differentiate between cancerous and benign lesions. The goal is to build a classification model for the malignant melanoma lesions based on the selected features.

Regarding the first phase, image acquisition, most of the existing approaches are suitable mainly for dermatoscopic or spectroscopic images and fail to provide a complete solution using images captured by smartphone cameras. Dermoscopic images are acquired under controlled clinical conditions by employing a liquid medium (or a nonpolarized light source) and magnifiers. This type of image includes features both on and below the skin surface, a property that is not yet achieved by standard cameras. Therefore, using dermatoscopic images as the basis of the skin cancer diagnosis app may limit its applicability to assisted-living scenarios. In contrast, the existing body of dermatological knowledge of melanoma diagnosis is almost completely based on dermatoscopic and spectroscopic images. This knowledge-transfer gap has been vastly ignored by the existing apps, resulting in unreliable extrapolation of results from these images to the ones captured by normal cameras.

To enable smartphone apps to capture dermatoscopic images, several mobile add-ons, such as DermLite (3Gen, San Juan Capistrano, California) and handyscope (FotoFinder Systems, Bad Bimbach, Germany), have been introduced to the market. Although these devices increase the usability and mobility of such technologies, the cost is too high to be practical for assisted-living apps.

For the second phase, common approaches employed for lesion segmentation are thresholding techniques (histogram

thresholding, adaptive thresholding, morphological thresholding, hybrid thresholding on optimal color channels, and differences in Gaussian filtering), deformable model techniques (iterative classification, deformable models, active contours, and adaptive snake and random walker algorithms), and clustering techniques (wavelet transform, wavelet neural networks, edge and region merging, and fuzzy sets) (reviewed in detail in [24]). Of these methods, histogram thresholding and region merging are most commonly used due to their simplicity and low computational costs, but they have a greater possibility of imperfect segmentation. Histogram thresholding methods use image histograms to determine one or more intensity values to separate pixels into groups, usually based on the Otsu lesion segmentation method [25]. While opting for a simpler, faster, and less expensive segmentation method, it is anticipated that the errors in segmentation will be resolved in the next phases. A robust segmentation method is required for smartphone-based imaging apps due to loosely controlled lighting and focal conditions.

For the third phase, the most commonly used features for lesion classification include color variation, border irregularity, asymmetry, texture, and shape [24]. These features are then fed to a classifier, such as SVM and KNN, to produce the final output.

As discussed in the “Introduction” section, there has been an increase in smartphone availability equipped with multi-core central processing units (CPUs), GPUs, rich multimedia touchscreens, and high-resolution image sensors. This allows people to become more proactive and involved in their own health-care processes. However, most of the systems working on mobile platforms (such as Lubax) have chosen to push all of the computation to the server side, using the mobile device

only for capturing, storing, and transmitting skin lesion images to a remote server [26]. As discussed for the wound assessment apps, this requires a stable data link and creates privacy and security challenges that are usually ignored by the developers, who assume that they will be handled by the (cloud) computing provider.

Nonetheless, a few systems perform the analysis of smartphone-captured or dermatoscopic images directly on a mobile device. For example, Wadhawan et al. [27] proposed an on-device library for melanoma detection based on a bag-of-features framework. They showed that the most computationally intensive and time-consuming algorithms (image segmentation and image classification) can achieve an accuracy and execution speed comparable to a desktop computer.

On-device processing of images captured by a smartphone has been reported in [28] and [23]. The system proposed in [28] uses a basic thresholding method for segmentation, followed by classification based on the standard color feature and border features, but the results have low accuracy (66.7%). In contrast, Doukas et al. [23] focused on the system integration, without describing in detail the features used to reach the reported accuracy of 77.06%.

In [7], a fully on-device melanoma diagnosis system is proposed that employs a combination of skin detection and edge localization for fast lesion segmentation. The segmented lesion is then used for feature extraction. The features used are motivated by the original features defined by dermatologists, also known as the ABCD feature set: asymmetry, border irregularity, color and texture variation, and diameter [7]. To cope with the limited resources available on the mobile device, the authors devised a new feature selection mechanism that takes into account the coordinate of the feature values to identify more discriminative features and efficiently process them. Given the feature set \mathcal{F} and the class label L , the feature selection problem is to find a set $\mathcal{S} \subset \mathcal{F}$ ($|\mathcal{S}| < |\mathcal{F}|$) such that it maximizes the relevance between L and \mathcal{S} , usually characterized in terms of MI. The hybrid criterion used for feature selection combines NMIFS and average neighborhood margin (ANM) maximization schemes in a single unified criterion defined by the following relation [7]:

$$U_{f_i} = \alpha \cdot Q_{f_i} + (1 - \alpha) \cdot \left[MI(L, f_i) - \frac{1}{|\mathcal{S}|} \sum_{f_j \in \mathcal{S}} NMI(f_i, f_j) \right], \quad (6)$$

where $|\cdot|$ is the cardinality of the set, Q_f is the quality of feature f (defined below), MI is the mutual information that measures the relevance between two discrete random variables X and Y (with alphabets \mathcal{X} and \mathcal{Y}), NMI is the normalized MI, and $\alpha \in [0, 1]$ is a weight factor that controls the influence of the ANM and MI in the proposed hybrid criterion.

The quality of the feature f (which is derived from ANM) is defined as

$$Q_f = \sum_{i=1}^n \left\| \left(\sum_{j=1}^{n_i^e} \frac{\|f_i - f_j\|_1}{n_i^e} - \sum_{j=1}^{n_i^o} \frac{\|f_i - f_j\|_1}{n_i^o} \right) \right\|_1, \quad (7)$$

where, for each sample i , n_i^e is the set of the most similar samples that are in the same class with sample i , and n_i^o is the set of the most similar samples that are not in the same class with i . The set of selected features is extracted and used by a classifier array to produce the detection results, reporting over 80% sensitivity and specificity.

A comparison of the existing smartphone apps (available for Android and iOS platforms) and proposed approaches is presented in Table 5 in terms of use of automated decision-making (i.e., machine learning), image modality, accuracy, test bed size, and on-device/on-server processing. As indicated by the reported accuracies, test bed size, and image modalities, smartphone-based skin cancer diagnosis apps are not developed enough to be used even by experts and, therefore, are not ready to be used for assisted-living scenarios.

Discussion

Smartphones are used in a wide range of situations in everyday life and thus provide a unique opportunity to help users daily. Offering several different sensors and computational power, smartphones have proved to be a promising tool for assisted living in both cities and isolated rural areas, with apps ranging from human-computer interfaces and augmented reality apps for disabled users to health-monitoring and fitness-tracking devices. Smartphone-based health-monitoring apps compose one of the main tracks of assisted-living apps, which have drawn increased attention due to the growth of the aging population as well as eHealth ecosystems. However, the challenge of developing a well-designed app for health monitoring should not be underestimated. To achieve the desirable accuracy, performance, and usability, techniques from different fields, such as interface design, sensor fusion, signal processing, image processing, and machine learning, should be properly combined to

Table 5. A comparison of representative existing systems used for skin cancer diagnosis.

	Machine Learning	Accuracy	Processing	Image Modality	Data Set	Operating System Platform
Abuzaghlleh et al. [26]	No	90%	Server	Dermoscopy	200	iOS
Wadhawan et al. [27]	Yes	79.26%	Local	Dermoscopy	347	iOS
Ramlakhan et al. [28]	Yes	66.7%	Local	Visible Light	83	iOS
Doukas et al. [23]	Yes	77.06%	Local/Server	Dermoscopy	3,000	Android
Do et al. [7]	Yes	84%	Local	Visible Light	184	Android
Lubax	Yes	Not Available	Local	Visible Light	>12,000	iOS/Android

function in near real-time and under computational, memory, and energy consumption constraints. We provide an overview of the goals, methodologies, and smartphone apps for health monitoring to validate whether different areas in this field have reached the level of maturity required for general public use, the target population.

Categorical assessment

At the very first step, it is important to determine the legacy algorithms and open-source codes, such as those from the sensor fusion, signal processing, image processing, and machine-learning fields. But as discussed in the “Mobile Image Processing Algorithms for Assisted Living” section, many of these algorithms cannot be used in for assisted-living apps unless they address the existing challenges, including computational, memory, and energy constraints. Algorithms such as segmentation, 3-D reconstruction, and machine learning should be used in many advanced assisted-living apps, but as they have the highest complexity and cost factors, they will require major changes before being incorporated into smartphone apps. There has been a movement toward migrating commonly used algorithms to smartphone platforms (e.g., the OpenCV package for smartphones, light-weight neural networks, and midlevel image classification algorithms), but more specific algorithms, such as specialized features, may still require significant adaptation for smartphone use. It is important to note that this compatibility issue is different from algorithm performances in terms of accuracy and robustness.

In addition to the underlying algorithms used, it appears that different categories of health-monitoring apps are at different development levels, most likely due to differences in the challenges faced. Lifestyle-tracking apps, such as pedometers and heart-rate-monitoring apps, can be readily used by the public because they address easier problems and, at the same time, can tolerate lower accuracies. For example, inconsistencies in heart rate readings (e.g., due to different finger placement or motion artifacts) do not substantially affect the usability of an app. The collected user information has no high-level privacy concerns and can be easily transmitted to and stored on a remote server using almost any available data link. Generally, these apps can be improved to provide better results in an uncontrolled environment, but, even at their current level, they can be used as a component for assisted living.

The challenges escalate as we move toward more advanced and critical apps, such as chronic wound assessment. These apps incorporate some level of artificial intelligence to help the user either gain some degree of expert capability (e.g., generic assessment of wound condition) or obtain an expert opinion without the need for an office visit (e.g., through image or statistics transmission). They perform a deeper analysis of the captured image or signal, extract relevant features, and then infer higher-level information through statistical analysis. Users are empowered with more advanced capabilities compared with the first category of assisted-living apps. To achieve this, however, the apps need to address the same problems as the first category (e.g., noise reduction and normalization) but

with a higher level of scrutiny. In addition, they face problems that are not present or are negligible for the first category, including estimating 2-D or even 3-D reconstruction, feature extraction in the presence of noise, and robust and customized classification using both local and contextual information. In the case of wound assessment, for example, the color and illumination normalization is itself an unresolved challenge. Obtaining basic and yet vital features such as wound dimensions can be obtained only by investing substantial computational, memory, and energy; external markers (of specific size and color) still need to be placed near the wound for camera calibration and illumination normalization.

These issues prevent many apps of this type from using more advanced features such as 3-D structure or video processing of the wound. In addition, contextual features [e.g., the periwound condition, exudation, and history information (e.g., history of diabetes)] that are routinely used by experts for wound grading have not been implemented in any existing system, regardless of platform. These challenges are the reason that the existing state of the art is used only by medical experts to assist them (and not replace them) in low-level tasks, e.g., wound measurements. Functionalities such as tissue classification have severe shortcomings that can mislead even experts. We conclude that many of the basic challenges facing this second category of health-monitoring apps and other assisted-living apps at the same level have been solved. However, these apps still have more advanced issues that need to be addressed before use by a larger target population.

Finally, the third and most advanced smartphone assisted-living app category targets problems that still exist without smartphone platform restrictions. It aims to address challenges that no definitive approach has yet resolved. The underlying problem for this category of apps is particularly sensitive and demands a solution with high accuracy. Due to the sensitivity (e.g., privacy or security issues) of this problem, it faces additional issues, such as a lack of proper and standardized test beds or availability of a robust and discriminative feature. Example of these problems are visible in the smartphone-based skin cancer diagnosis systems, reviewed in the “Chronic Wound Assessment” section. An acceptable skin cancer diagnosis system requires near-expert accuracy to avoid endangering the user or misleading the expert. To reach that level of accuracy, different modalities of data should be combined and evaluated, and the system should be extensively validated on medically proven cases. But our review shows the contrary. There is a significant variation in the quality, modality, and number of images used in different proposed works. The numbers of images range from fewer than 100 to more than 12,000, making the results incomparable. Moreover, many apps use dermatoscopic images in their process when they are targeted for smartphone. This divergence without visible and verified model transfer is one of the limitations that undermine the existing works.

In addition, processes used for preprocessing, feature extraction, and classification are not still accurate enough to address the problem at hand. There seem to be no substantial color and

illumination normalization steps, and the selected features fall short of providing medically reliable information about the skin lesion. There is no 3-D structure analysis of the lesion, and the classifiers used seemed to be too generic and, therefore, unable to distinguish between cancerous and benign lesions. Although not assessed in the literature, these approaches would logically fail in the face more challenging conditions, including variations in skin type and color (proven to be challenging even for advanced image processing approaches such as face recognition/detection), illumination changes and cast shadows (causing color and 2-D/3-D model distortions), or imperfect image captures such as with motion artifacts.

Crippled by these challenges, it is understandable that apps such as skin cancer diagnosis have not evolved to use higher-level information in the clinical field. For example, these systems cannot distinguish between slow- and fast-growing melanomas, which require different treatments. It is worth mentioning that this may be a cyclic problem, meaning that, with a lack of a sufficient number of proper samples from cancerous and benign lesions, offline feature selection is likely to fail in finding a powerful and robust feature set. Propagating this problem to the classification step, this all results in a low-performing system. We therefore conclude that this type of app still needs to fill several gaps even to be used by experts.

A possible solution may be to first address the problems of this sort in a more relaxed setting, e.g., with no computational or memory constraints and in a more controlled data acquisition scenario. Upon finding a more relaxed solution, the next step could be adapting the techniques to more constrained settings for expert user population. Finally, release for general public use as the ultimate goal has to evolve with caution to reduce the risk of possible damages. For example, in the case of the melanoma diagnosis app, even given a medical-grade accuracy of diagnosis, the psychological aspects should be taken into account to prevent misleading users [29]. An example of the psychological effects of this app is illustrated by comparing the output to be “cancerous/benign” with “Please check with dermatologist/please check again in one month.” The former output gives an impression of severity and has related psychological risks (e.g., depression) and, in the case of false-positive and false-negative results, may dangerously change the user’s behavior (e.g., dismissing any future detection results or avoiding medical attention when needed). In contrast, the latter output indicates a possibility of an error in the results (both regarding positive and negative results) and may reduce the chance of undesired behavioral changes in the user.

Open challenges

On the basis of our review on the three categories of smartphone apps, we summarize the challenges to be addressed in future works.

- *Lack of clinical-based evidence.* A predominant characteristic of many existing health-monitoring apps is no or minimal involvement of clinical experts in the product design/evaluation phases. This issue is critical because these apps

target nonexpert users, so catastrophic results are a possibility. For example, the Federal Trade Commission acts against apps that pretend to be medical devices capable of diagnosing melanoma [33]. As discussed previously, many of these apps currently fail to provide assistance to field experts, not to mention the general public.

- *Prevalence of on-server processing.* Although current smartphones are equipped with powerful processing units and multimodal high-resolution sensors, many of the apps rely on on-server processing, limiting the usage to the availability of a data link. Using the smartphone capabilities to at least partially process the data not only reduces the data link reliability requirement but also may help remove sensitive data without losing information by processing raw data into feature sets.
- *Video data platforms.* Using video data instead of a single 2-D image for the medical apps is still undeveloped due to the computational overhead. A video can certainly offer more rich data than a single image, and few medical apps venture in this direction, e.g., heart rate estimation. The immediate consequence of using a video is that the 3-D representation of a scene can be computed, and more interactive, user-friendly apps can be created. For instance, the 3-D morphology can be obtained, and the associated measurements of skin lesions (e.g., wounds or moles) can be derived and explored by physicians. Compared with the 2-D model, we believe that adding a new dimension to the visualized data will provide clinicians the ability to diagnose early and better treat ill patients.
- *3-D vision enabling.* Whether a small-sized (e.g., a skin mole) or large-sized (e.g., a building) structure, reconstructing the 3-D structure of a target scene may reveal valuable information that can dramatically improve the performance of assisted-living apps. The state of the art of mobile-based 3-D reconstruction has several limitations in terms of resolution (particularly regarding small structures such as wounds or skin moles); high computational, memory, and energy costs (particularly for passive approaches); and requirements for additional hardware and markers for acquisition and calibration. These have prevented many smartphone apps from gaining practical benefits from 3-D reconstruction approaches. Therefore, improvements are needed in both the algorithms and hardware issues of 3-D scene reconstruction. For example, light-weight but accurate algorithms for finding point correspondence and specialized optimization solutions and built-in 3-D acquisition or even on-chip and real-time 3-D reconstruction modules can enable smartphone apps to use the information embedded in 3-D, without suffering from current consequences.
- *Multimodality data processing and fusion.* The goal of many of the assisted-living apps, particularly the ones for advanced health monitoring, seems to be achievable only by processing and fusion of multimodal data that can be obtained by the smartphone (e.g., images, videos, locations, motion, etc.) and also through other contextual information databases

(e.g., the user's family and history records, local and global trends, etc.). The next step is investigating the best approaches for integrating these different data modalities.

- *Health-care ecosystem.* Another facet that needs development in the assisted-living and, in particular, health-monitoring apps is how they can integrate with the larger health-care ecosystem. This integration can provide new information regarding the app goal and can also share the app-specific data with the ecosystem. With this two-way integration, app functionalities can be personalized and patient-oriented, surpassing the limitation of any single assisted-living app.
- *Big-data- and cloud-based services.* One of the steps in making multimodal data available for assisted-living smartphone apps involves cloud-based solutions. Proper integration of smartphone apps with cloud-based services, such as Amazon Web Services, IBM Bluemix, and Microsoft Azure, will provide access to the required wide range of multimodal data and provide insights based on the entire collection of data gathered from different devices and different apps using big data analytics approaches. It is worth noting that big-data- and cloud-based solutions are different from health-care ecosystem integrations because they make possible a meta-analysis of the entire data plane by smartphone apps, which is otherwise impossible to obtain. In this sense, big data analytics is also different from generic on-server processing. In fact, a new trend of assisted-living apps can be built on these services to provide “X as a Service” models (e.g., telehealth or dermatology as a service). In addition, the use of these services will also add to the complexity of the system and therefore bring forth more challenges, such as the efficiency of meta-analysis incorporation, the nature of data to be sent to and received from these services, and app usability in the case of unreliable or missing data links.
- *Security and privacy.* With the distribution of data over different components in cloud-based services and health-care ecosystem comes the important issue of data privacy and security. Without appropriate security mechanisms in place, the transmitted information is susceptible to attacks, such as eavesdropping, denial of service, and man-in-the-middle attacks. Most of the assisted-living apps have assumed that these security measures are already in place and, therefore, neglected the required aspects of transmitted data, leaving possible security holes in the system. Although many of the security aspects may be handled by third-party service providers, issues such as identity-specific contents (e.g., images or history records) are too app-specific to be covered. Therefore, there is a pressing requirement for app-specific research on security and privacy problems and solutions in assisted-living and, particularly, health-monitoring apps [30].

Conclusions

It is difficult to review all possible apps and models for assisted-living scenarios and health monitoring. To obtain a broad view

of the current state of the art, we therefore reviewed representatives of each important category of smartphone-based health monitoring, with an eye on the applicability of the derived insights to other assisted-living apps. We reviewed commonly used algorithms and their capabilities and corresponding overheads and then image-based heart-rate-monitoring apps as a representative of low-risk assisted-living apps that are ready for public use. In the second category, we reviewed wound assessment and monitoring apps, which represent assisted-living apps with midlevel risk, provide basic functionalities, and may be used to assist experts but are not ready for public use. Finally, we reviewed a smartphone-based skin cancer diagnosis app in the most advanced and least developed category. This app represents high-risk assisted-living apps with the lowest level of development that are not yet capable of providing assistance even to experts.

On the basis of our reviews and discussions, we conclude that, although the abilities of smartphone apps to address problems in assisted living are improving, there are still major challenges to overcome before they become an everyday, personalized health-care provision tool in the medical field and in a clinical practice. The current apps lack a solid evidence-based approach and assurance of the credibility of the content. The challenges of smartphone-based assisted-living apps must be resolved by a joint effort of different organizations, including app developers, signal processing and machine-learning R&D groups, cloud-based service providers, and hardware providers. Efforts of this kind have already begun, with the OpenCV package for smartphones, the IBM Bluemix application program interface for smartphones apps, and Microsoft Windows 10 device independence. App developers can then use the solutions to both improve their apps and stimulate growth in such efforts. Although prevalent problems have prevented public use of many assisted-living apps, the trend of advanced smartphone app development by both small start-ups and multinational corporations suggests a bright future for smartphone assisted-living apps.

Acknowledgments

Hossein Nejati and Victor Pomponiu contributed equally for this article. The material reported in this article is supported by the Massachusetts Institute of Technology–Singapore University of Technology and Design (SUTD) International Design Centre (IDC). Any findings, conclusions, recommendations, or opinions expressed are ours and do not necessarily reflect the views of the IDC.

Authors

Hossein Nejati (hossein_nejati@sutd.edu.sg) received his B.S. degree in computer science and engineering from Shiraz University, Iran, and his Ph.D. degree from the National University of Singapore. He is a postdoctoral fellow at the Singapore University of Technology and Design. His research interests include medical image processing, biosignal data mining, brain image analysis, signal processing, machine learning, and data visualization of medical images and biosignals.

Victor Pomponiu (victor_pomponiu@sutd.edu.sg) received his B.S. and M.S. degrees in computer science from the Polytechnic University of Bucharest, Romania, in 2007 and 2009, respectively, and his Ph.D. degree in computer science from the Università degli Studi di Torino, Italy, in 2012. He is a postdoctoral fellow at the Singapore University of Technology and Design. His research interests include multimedia security and medical image analysis (in particular, image-based diagnosis of breast and skin cancer and wound assessment research).

Thanh-Toan Do (thanhtoan_do@sutd.edu.sg) received his Ph.D. degree in computer science from the University of Rennes 1, France, in 2012. He is a postdoctoral researcher at the Singapore University of Technology and Design. His research interests include computer vision, multimedia, and forensics.

Yiren Zhou (yiren_zhou@myemail.sutd.edu.sg) received his B.S. degree from the University of Science and Technology of China in 2013. He is a Ph.D. student in information systems technology and design at the Singapore University of Technology and Design. His research interests include image processing, computer vision, and machine learning.

Sahar Iravani (sahariravani.ee@gmail.com) received her B.S. degree in electrical engineering from Tabriz University, Iran, and her M.S. degree with honors in electrical engineering from the Babol Noshirvani University of Technology, Iran, in 2012 and 2015, respectively. She is currently with the Singapore University of Technology and Design. Her research interests include machine learning, deep neural networks, and biomedical image processing.

Ngai-Man Cheung (ngaiman_cheung@sutd.edu.sg) received his B.Eng. degree in computer engineering from the University of Hong Kong and his Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles. He is an assistant professor at the Singapore University of Technology and Design. His research interests include image and signal processing with application to health care.

References

- [1] M. N. K. Boulos, S. Wheeler, C. Tavares, and R. Jones, "How smartphones are changing the face of mobile and participatory healthcare: An overview, with example from ecaalyx," *Biomed. Eng. Online*, vol. 10, no. 24, 2011. DOI: 10.1186/1475-925X-10-24
- [2] N.-M. Cheung and D. Koh. (2014, Jan.). Mobile imaging system for early diagnosis of skin cancer. *IEEE Life Sci. Newsl.* [Online]. Available: <http://lifesciences.ieee.org/publications/newsletter/january-2014/482-mobile-imaging-system-for-early-diagnosis-of-skin-cancer>
- [3] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Express*, vol. 18, no. 10, pp. 10762–10774, May 2010.
- [4] S. Kwon, H. Kim, and K. S. Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone," in *Proc. 2012 Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, San Diego, CA, pp. 2174–2177.
- [5] F.-C. Huang, G. Wetzstein, B. A. Barsky, and R. Raskar, "Eyeglasses-free display: Towards correcting visual aberrations with computational light field displays," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 59:1–59:12, July 2014.
- [6] A. Börve, K. Terstappen, C. Sandberg, and J. Paoli, "Mobile teledermoscopy—there's an app for that!," *Dermatol. Pract. Concept.*, vol. 3, no. 2, pp. 41–48, Apr. 2013.
- [7] T.-T. Do, Y. Zhou, H. Zheng, N.-M. Cheung, and D. Koh, "Early melanoma diagnosis with mobile imaging," in *Proc. 2014 36th Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, Chicago, IL, pp. 6752–6757.
- [8] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

- [9] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Computer Vis.*, vol. 59, no. 2, pp. 167–181, Sept. 2004.
- [10] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," *IEEE Trans. Neural Networks*, vol. 20, no. 2, pp. 189–201, Feb. 2009.
- [11] Q. Wang, Z. Yu, C. Rasmussen, and J. Yu, "Stereo vision-based depth of field rendering on a mobile device," *J. Electron. Imag.*, vol. 23, no. 2, p. 023009, 2014.
- [12] H. Chikatsu and Y. Takahashi, "Comparative evaluation of consumer grade cameras and mobile phone cameras for close range photogrammetry," *Proc. SPIE*, vol. 7447, Sept. 2009.
- [13] *IEEE Standard for Camera Phone Image Quality*, IEEE Standard P1858-2012.
- [14] S. Mayu, M. Teruhiro, and Keiichi, Y. "Estimating heart rate variation during walking with smartphone," in *Proc. 2013 ACM Int. Joint Conf. Pervasive and Ubiquitous Computing*, New York, pp. 245–254.
- [15] G. Gaoan and Z. Zhenmin, "Heart rate measurement via smart phone acceleration sensor," in *Proc. 2014 Int. Conf. Smart Computing (SMARTCOMP)*, Hong Kong, pp. 295–300.
- [16] A. Pal, A. Sinha, A. D. Choudhury, T. Chattopadhyay, and A. Visvanathan, "A robust heart rate detection using smart-phone video," in *Proc. Third ACM MobiHoc Workshop on Pervasive Wireless Healthcare*, New York, 2013, pp. 43–48.
- [17] S. Xu, L. Sun, and G. K. Rohde, "Robust efficient estimation of heart rate pulse from video," *Biomed. Opt. Express*, vol. 5, no. 4, pp. 1124–1135, Apr. 2014.
- [18] Y. Sun, S. Hu, V. Azorin-Peris, S. Greenwald, J. Chambers, and Y. Zhu, "Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise," *J. Biomed. Opt.*, vol. 16, no. 7, p. 077010, 2011.
- [19] B. M. Bates-Jensen, "Chronic wound assessment," *Nurs. Clin. North Am.*, vol. 34, no. 4, pp. 799–845, 1999.
- [20] N. D. J. Hettiarachchi, R. B. H. Mahindaratne, G. D. C. Mendis, H. T. Nanayakkara, and N. D. Nanayakkara, "Mobile based wound measurement," in *Proc. 2013 IEEE Point-of-Care Healthcare Technologies (PHT)*, Bangalore, India, 2013, pp. 298–301.
- [21] M. A. L. Wang, P. C. Pedersen, D. M. Strong, B. Tulu, E. Agu, and R. Ignatz, "Smartphone-based wound assessment system for patients with diabetes," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 477–488, Feb. 2015.
- [22] H. Wamou, Y. Lucas, and S. Treuillet, "Enhanced assessment of the wound-healing process by accurate multiview tissue classification," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, pp. 315–326, Feb. 2011.
- [23] C. Doukas, P. Stagkopoulou, C. T. Kiranoudis, and I. Maglogiannis, "Automated skin lesion assessment using mobile technologies and cloud platforms," in *Proc. 2012 Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, San Diego, CA, pp. 2444–2447.
- [24] M. Sadeghi, T. K. Lee, D. McLean, H. Lui, and M. S. Atkins, "Detection and analysis of irregular streaks in dermoscopic images of skin lesions," *IEEE Trans. Med. Imag.*, vol. 32, no. 5, pp. 849–861, May 2013.
- [25] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285–296, pp. 23–27, 1975.
- [26] O. Abuzagheh, B. D. Barkana, and M. Faezipour, "Noninvasive real-time automated skin lesion analysis system for melanoma early detection and prevention," *IEEE J. Trans. Eng. Health Med.*, vol. 3, pp. 1–12, Apr. 2015.
- [27] T. Wadhawan, N. Situ, H. Rui, K. Lancaster, X. Yuan, and G. Zouridakis, "Implementation of the 7-point checklist for melanoma detection on smart handheld devices," in *Proc. 2011 Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, Boston, pp. 3180–3183.
- [28] K. Ramlakhan and Y. Shang, "A mobile automated skin lesion classification system," in *Proc. 2011 IEEE 23rd Int. Conf. Tools with Artificial Intelligence*, Boca Raton, FL, pp. 138–141.
- [29] E. D. Boudreaux, M. E. Waring, R. B. Hayes, R. S. Sadasivam, S. Mullen, and S. Pagoto, "Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations," *Transl. Behav. Med.*, vol. 4, no. 4, pp. 363–371, Dec. 2014.
- [30] Food and Drug Administration. (2013, Sept. 25). Mobile medical applications: Guidance for industry and food and drug administration staff. [Online]. Available: <http://www.fda.gov/downloads/MedicalDevices/UCM263366.pdf>
- [31] P. Moulon. (2016). OpenMVG: Open multiple view geometry. [Online]. Available: imagine.enpc.fr/~moulonp/openMVG
- [32] Movidius. (2016, Jan. 27). Google and Movidius to enhance deep learning capabilities in next-gen devices. [Online]. Available: <http://www.movidius.com/news/google-and-movidius-to-enhance-deep-learning-capabilities-in-next-gen-devices>
- [33] Federal Trade Commission. (2015, Feb. 23). FTC cracks down on marketers of "melanoma detection" apps. [Online]. Available: www.ftc.gov/news-events/press-releases/2015/02/ftc-cracks-down-marketers-melanoma-detection-apps

Vishal M. Patel, Rama Chellappa,
Deepak Chandra, and Brandon Barbello

Continuous User Authentication on Mobile Devices

Recent progress and remaining challenges

Recent developments in sensing and communication technologies have led to an explosion in the use of mobile devices such as smartphones and tablets. With the increase in the use of mobile devices, users must constantly worry about security and privacy, as the loss of a mobile device could compromise personal information. To deal with this problem, continuous authentication systems (also known as *active authentication systems*) have been proposed, in which users are continuously monitored after initial access to the mobile device. In this article, we provide an overview of different continuous authentication methods on mobile devices. We discuss the merits and drawbacks of the available approaches and identify promising avenues of research in this rapidly evolving field.

Introduction

Traditional methods for authenticating users on mobile devices are based on explicit authentication mechanisms such as a password, a personal identification number (PIN), or a secret pattern. Studies have shown that users often choose a simple, easily guessed password like “12345,” “abc1234,” or even “password” to protect their data [1], [2]. As a result, hackers could easily break into many accounts just by trying the most commonly used passwords. Also, when secret patterns are used for gaining initial access on the mobile devices, users tend to use the same pattern over and over again. As a result, they leave oily residues or smudges on the screen of the mobile device. It has been shown that with special lighting and high-resolution photography, one can easily deduce the secret pattern (see Figure 1) [3].

Digital Object Identifier 10.1109/MSP.2016.2555335
Date of publication: 1 July 2016





FIGURE 1. A smudge attack [3]. A user's secret pattern can be determined with special lighting and high-resolution camera.

Furthermore, recent studies have shown that about 34% of users did not use any form of authentication mechanism on their devices [4]–[7]. In these studies, inconvenience was cited as one of the main reasons that users did not employ any authentication mechanism on their devices [6], [7]. The authors of [7] demonstrated that mobile-device users considered unlock screens unnecessary in 24% of situations and that they spent up to 9% of their smartphone use time unlocking the screen. Furthermore, as long as the mobile phone remains active, typical devices incorporate no mechanisms to verify that the user originally authenticated is still the user in control of

the device. Thus, unauthorized individuals may improperly obtain access to a user's personal information if a password is compromised or if the user does not exercise adequate vigilance after initial authentication.

To overcome these issues, both the biometrics and security research communities have developed techniques for continuous authentication on mobile devices. These methods essentially make use of physiological and behavioral biometrics, using built-in sensors and accessories such as the gyroscope, touch screen, accelerometer, orientation sensor, and pressure sensor, to continuously monitor user identity. For instance, physiological biometrics such as those of the face can be captured using the front-facing camera of a mobile device and used to continuously authenticate a mobile-device user. On the other hand, sensors such as the gyroscope, touch screen, and accelerometer can be used to measure behavioral biometric traits such as gait, touch gestures, and hand movement. Figure 2 highlights some of the sensors and accessories available in a modern mobile device. These sensors are capable of providing raw data with high precision and accuracy, and are useful in monitoring three-dimensional (3-D) device movement or positioning or to monitor changes in the ambient environment near a mobile device. Note that the terms *continuous authentication*, *active authentication* [8], *implicit authentication* [9], [10], and *transparent authentication* [11] have been used interchangeably in the literature.

Our goal in this article is to survey recent developments in continuous authentication, discuss their advantages and limitations, and identify areas still open for exploration. Development of feasible and robust continuous authentication systems for

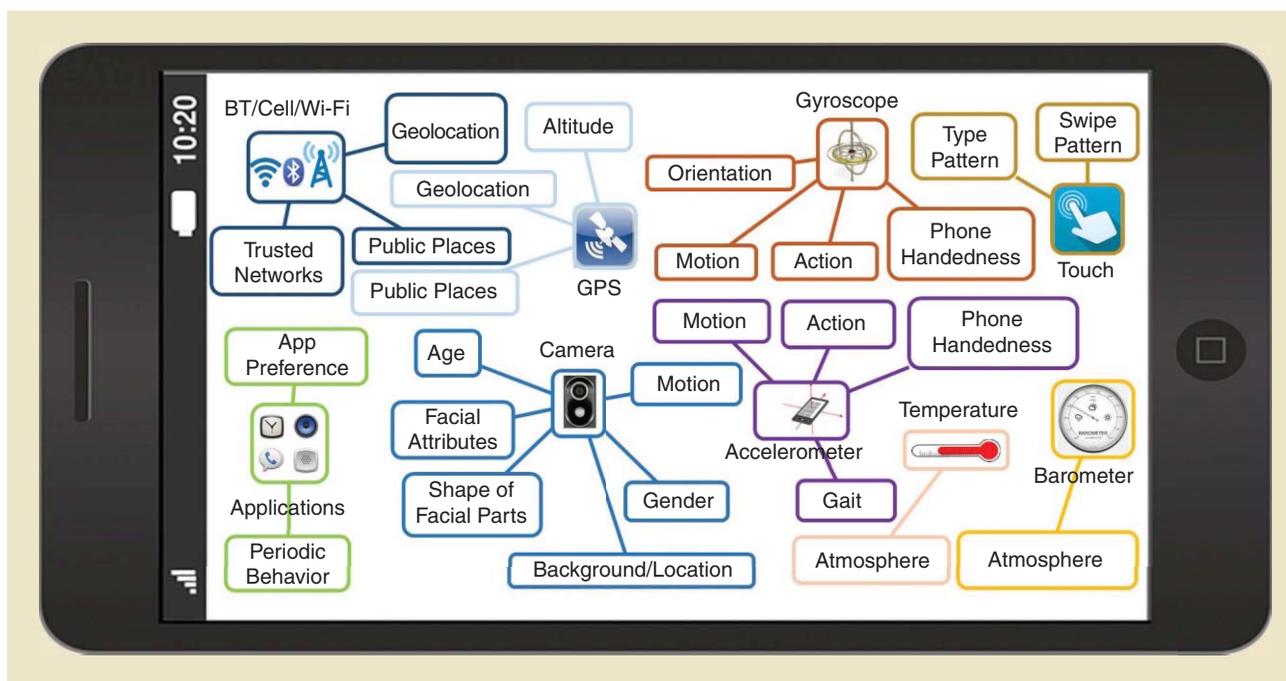


FIGURE 2. Sensors and accessories available in a mobile device. Raw information collected by these sensors can be used to continuously authenticate a mobile-device user. GPS: global positioning system.

mobile devices is important, as we are becoming increasingly dependent on mobile devices.

Continuous authentication approaches

Figure 3 shows the basic concept of a biometrics-based mobile device continuous authentication system [12]. Biometric modalities such as gait, face, keystroke, or voice are measured by the sensors and accessories that are in a mobile device. Then, the biometric system will determine whether these biometric traits correspond to a legitimate user or not. If the features do correspond to a legitimate user, the biometric system will continue processing the new incoming data. However, if the biometric system produces a negative response, the system will ask the user to verify his or her identity by using the traditional explicit authentication methods based on PIN, face, or secret pattern. If the user is able to provide identity verification, the mobile device will continue in service; otherwise, it will be locked.

In a practical continuous authentication system, the entire process happens in real time. A plethora of mobile continuous authentication methods have been proposed in the literature. In what follows, we review a few recent methods based on physiological as well as behavioral biometrics for continuous authentication.

Touch dynamics

Touch dynamics is one of the most commonly used continuous authentication methods for mobile devices. In touch dynamics, touch screen input is used as a data source. In particular, screen touch gestures—the way users swipe their fingers on the touch screen of their mobile devices—are used as a behavioral biometric to continuously authenticate users while they perform basic smartphone operations. In these methods, a behavioral feature vector is extracted from the recorded screen touch data, and a discriminative classifier is trained on these extracted features for authentication. Figure 4 shows some swipes performed by eight different users while reading text on an Android device [13]. It is interesting to see that even for the same task, touch data of different users show significant differences. In addition to the x and y coordinates of each swipe, information such

as finger pressure, the screen area covered by each finger, and time information can be used to extract useful features.

A swipe or a stroke on the touch screen is a sequence of touch data when the finger is in touch with the screen of the mobile device. Every swipe s can be encoded as a sequence of vectors

$$s_i = (x_i, y_i, t_i, p_i, A_i, O_i^f, O_i^{ph}), \quad i = \{1, 2, L, N\}, \quad (1)$$

where x_i, y_i are the location points, and t_i, p_i, A_i, O_i^f , and O_i^{ph} are the time stamp, the pressure on screen, the area occluded by the finger, the orientation of the finger, and the orientation of the phone (landscape or portrait), respectively. Here, N is the total number of swipes. Based on these measurements, a 30-dimensional feature vector was proposed in [13] for each swipe: midstroke area covered; 20% pairwise velocity; midstroke pressure; direction of end-to-end line; stop x ; start x ; average direction; start y ; average velocity; stop y ; stroke duration; direct end-to-end distance; length of trajectory; 80% pairwise velocity; median velocity at last three points; 50% pairwise velocity; 20% pairwise acceleration; ratio of end-to-end distance and length of trajectory; largest deviation from end-to-end line; 80% pairwise acceleration; mean resultant length; median acceleration at first five points; 50% deviation from end-to-end line; interstroke time; 80% deviation from end-to-end line; 20% deviation from end-to-end line; 50% pairwise acceleration; phone orientation; midstroke finger orientation; and up/down/left/right flag. After feature analysis, three of these features were discarded and the remaining 27 were evaluated using a kernel support vector machine (SVM) and k -nearest-neighbors (k NNs) classifiers on a data set consisting of 41 users' touch gestures. It was shown that these classifiers can achieve equal error rates (EERs) between 0% and 4%, depending on the application scenario [13]. Similar features have been used in [14]–[16] for touch gesture-based continuous authentication. For classification, nonlinear sparse representation-based classifiers were used in [16], while ten different classification algorithms were evaluated in [15].

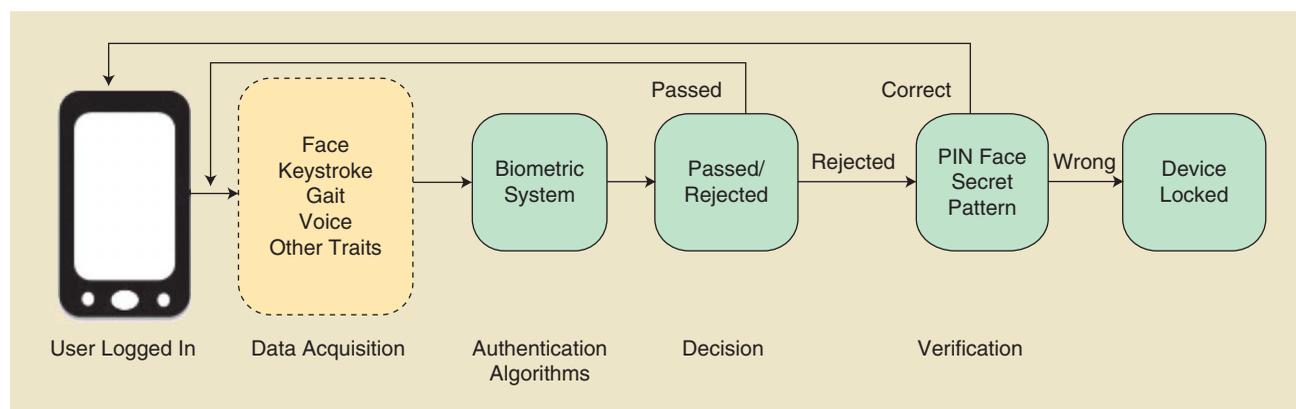


FIGURE 3. A biometrics-based mobile continuous authentication framework [12].



FIGURE 4. Swipes of eight different users while reading text [13]. Different colors are used to show different users' swipes.

The methods presented in [13]–[16] are essentially based on the fact that only a single finger is in contact with the touch screen while users are performing basic operations. In practice, many applications require users to employ two or more fingers to perform a particular task, such as zooming in and zooming out by pinching and spreading two fingers. More general multitouch, gesture-based continuous authentication systems have also been proposed in the literature [17], [18]. Similar to single-finger gestures, in [17] x and y coordinates, direction of finger motion, finger motion speed, pressure at each sampled touch point, and the distance between multitouch points are used to extract multitouch gesture features. On the other hand, in [18] a second-order autoregressive model is used for modeling multitouch sequences, and a mutual information-based metric is used for multitouch gesture recognition.

Different from the touch gesture features discussed above, an image-based feature called graphic touch gesture feature (GTGF) was proposed in [19] for modeling touch dynamics. In this approach, swipe geometry traits are converted to the image space so the dynamics of swipes can be explicitly modeled. Furthermore, the pressure dynamics are emphasized by fusing them with the movement dynamics. This method was later extended in [20] by building a touch gesture appearance model from the GTGF. The model learns the intraperson variations of the GTGF in the image domain by statistical analysis and is capable of synthesizing new instances according to a new probe swipe. Furthermore, these methods are applicable to both single-finger

and multifinger swipes. Figure 5 shows the GTGF features extracted from two users.

Table 1 compares all the aforementioned touch dynamics-based continuous authentication methods. Here, the false accept rate (FAR) and the false reject rate (FRR) are presented for one of the studies. As can be seen from this table, some methods achieve very low EER values on certain data sets. These studies have demonstrated that touch gestures can be used as a promising behavioral biometric for continuous user authentication of mobile devices.

Face recognition

Another continuous authentication system that is widely used for continuously monitoring a user's identity on a mobile device is based on face recognition. A generic face recognition system consists of three main stages. In the first, faces are detected from the images or videos captured by smartphones' front-facing cameras. Then, holistic or local features are extracted from the detected faces. Finally, these features are passed on to a classifier for authentication. A number of different methods have been proposed in the literature for detecting and recognizing faces on mobile devices. In what follows, we briefly review some of these methods.

In [21], the feasibility of face and eye detection on cell phones was evaluated using the Adaboost cascade classifiers with Haar-like and local binary pattern (LBP) features [22], [23] as well as a skin color-based detector. On a Nokia N90 mobile phone that has an ARM9 220-MHz processor

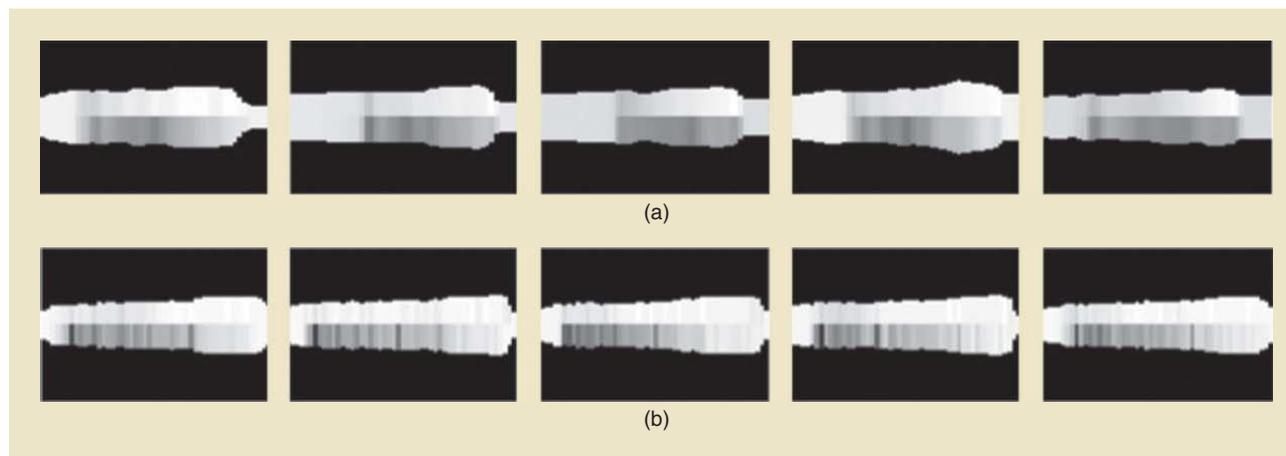


FIGURE 5. The GTGF features corresponding to two different users [19]: (a) shows the GTGF features corresponding to five touch gestures of a single user, while (b) shows the GTGF features extracted from five swipes of a different user.

and a built-in memory of 31 MB, the researchers reported that the Haar + Adaboost method can detect faces in 0.5 seconds from 320×240 images. This approach, however, is not effective when wide variations in pose and illumination are present or when the images contain partial or clipped images. To deal with these issues, a deep convolutional neural network (DCNN)-based method was recently developed in [24] for detecting faces on mobile platforms. In this method, deep features are first extracted using the first five layers of Alexnet [25]. Different-size sliding windows are considered, to account for faces of different sizes, and an SVM is trained for each window size to detect faces of that particular size. Then detections from all the SVMs are pooled together, and some candidates are suppressed based on an overlap criterion. Finally, a single bounding box is output by the detector. It was shown that this detector is quite robust to illumination change and is able to detect partial or extremely posed faces. A few sample positive detections from the University of Maryland–Active Authentication (UMD-AA) data set [26] are shown in Figure 6. The DCNN-based detections are marked in red, while the ground truth is in yellow. Another part-based method for detecting partial and occluded faces on mobile devices was developed in [27]. This method is based on detecting facial segments in the given frame and clustering them to obtain the region that is most likely to be a face.

In terms of face recognition on mobile devices, a method based on a one-class SVM was proposed in [28]. In

this approach, faces are detected using the Viola–Jones detector [22]. Histogram equalization is then applied on the detected images to normalize the effect of illumination. Finally, bidimensional Fourier transform features are extracted from the normalized images and fed into a one-class SVM for authentication. In addition to developing face- and eye-detection methods on mobile devices, [21] also developed a method for face recognition based on LBP features. It was shown that their proposed continuous face authentication system, including face detection and recognition, can process about two frames per second on a Nokia N90 mobile phone with an ARM9 processor with 220 MHz. Average authentication rates of 82% and 96% for images of size 40×40 and 80×80 , respectively, were reported in [21]. In [26], several face recognition methods were evaluated on a data set of 750 videos from 50 users collected over three sessions with different illumination conditions. A face-based continuous authentication method was recently developed in [12] that uses gyroscope, accelerometer, and magnetometer data to correct for camera orientation and the orientation of the face image. In [29], a sensor-assisted mobile face recognition system was proposed that utilizes motion and light sensors to defend against media and virtual camera attacks.

Visual attributes are essentially labels that can be given to an image to describe its appearance [30]. A facial attribute-based continuous authentication method was recently proposed in [31]. Figure 7 gives an overview of this method. Given

Table 1. Key touch dynamics-based continuous authentication methods. The best results from the corresponding papers are reported.

Study	Number of Users	Classifiers	Feature Dimension	Performance (%)
Frank et al. [13]	41	SVM, kNN	27	EER: 0.00–4.00
Zhang et al. [16]	50	Sparsity-based classifiers	27	EER: 0.77
Li et al. [14]	75	SVM	10	EER: ~ 3.0
Feng et al. [17]	40	Random forest, J48 tree, Bayes' net	53	FAR: ~ 7.50, FRR: ~ 8.00
Serwadda et al. [15]	138	Ten different classifiers	28	EER: 10.50
Zhao et al. [20]	78	L_1 distance	100×150 image	EER: 6.33–15.40

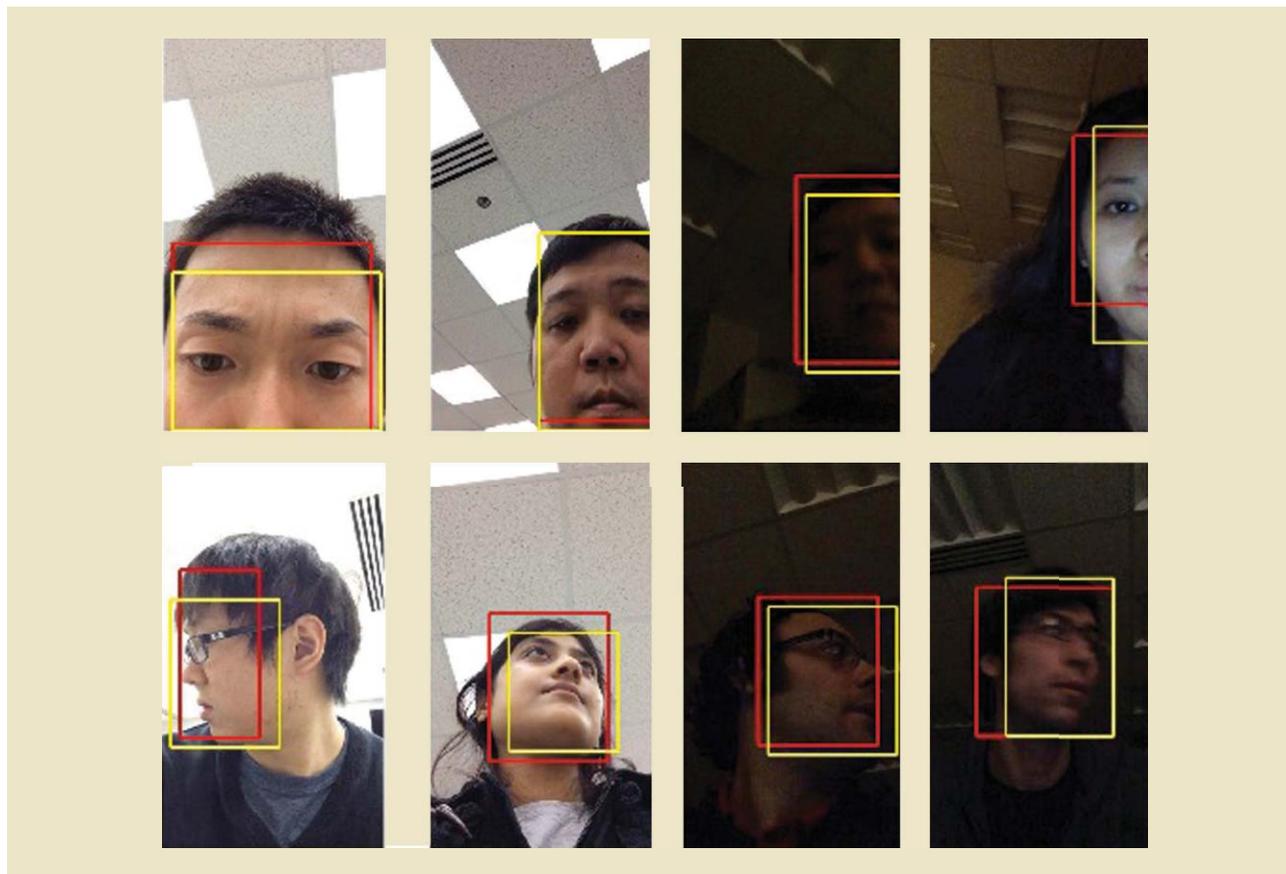


FIGURE 6. Examples of positive detections with pose variations and occlusion on the UMD-AA data set. The detector's output is in red, while ground truth is in yellow [24].

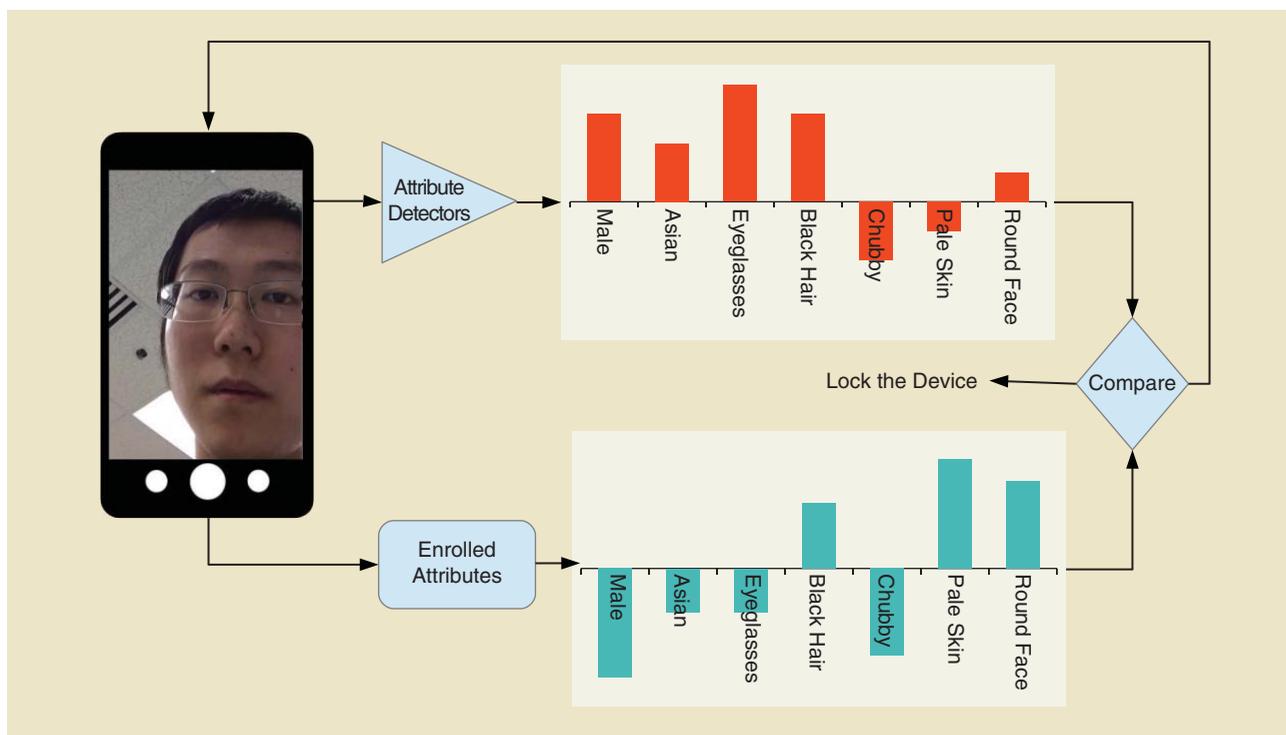


FIGURE 7. An overview of the attribute-based authentication method proposed in [31].

a face image sensed by the front-facing camera, pretrained attribute classifiers are used to extract a 44-dimensional attribute feature. The binary attribute classifiers are trained using the PubFig data set [30] and provide compact visual descriptions of faces. The score is determined by comparing the extracted attribute features with the features corresponding to the enrolled user. These score values are essentially used to continuously authenticate a mobile-device user. Furthermore, it was shown that the attribute-based method can be fused with an LBP-based method such as in [21] to obtain improved matching performance.

Table 2 summarizes key face-based continuous authentication methods. Here, the recognition rate (RR), true accept rate (TAR), and average authentication rate (AAR) are noted for some of the studies.

Gait dynamics

Gait dynamics-based continuous authentication systems identify users based on how they walk. The data needed for gait-based authentication are often measured by the built-in accelerometer and gyroscope sensors. Once the raw data are measured, discriminative features are extracted, which are then fed into a classifier to distinguish users. In recent years, several methods have been developed for gait-based recognition on mobile devices [32]–[38]. These methods differ essentially in the types of features extracted from the raw data for classification or the types of classification methods used for authentication. For instance, methods based on correlation, frequency domain analysis, and data distribution statistics are used in [32], while methods based on dynamic time warping are used in [36] and [37]. Rather than using the gait cycles for extracting features, [35] proposes an application of hidden Markov models (HMMs) for gait recognition. In particular, a sensor orientation invariant gait representation called *gait dynamic images (GDIs)* was proposed in [39]. Given a 3-D time series captured by a three-axis accelerometer, its GDI is calculated by the cosine similarity of the motion measurement at time t with the time-lagged signal of lag l . Figure 8 shows an example of raw three-axis accelerometer data and their

Table 2. A summary of key face-based continuous authentication methods.

Study	Number of Users	Method/Features	Performance (%)
Abeni et al. [28]	32	1-class SVM/Fourier transform	EER: 3.95–7.92
Hadid et al. [21]	12	Histogram intersection distance/LBP	AAR: 82–96
Fathy et al. [26]	50	Nine different classifiers/MEEN	RR: ~95
Crouse et al. [12]	10	SVM/biologically inspired model	TAR: ~40–50 @FAR 0.1
Samangouei et al. [31]	50	Attributes	EER: 13–30

MEEN: mouth, left eye, right eye, nose.

corresponding GDI. As can be seen from this figure, since GDI is invariant in regard to sensor orientation, it shows much better consistency before and after sensor rotation. Also, pace-independent gait recognition approaches have been proposed in [34] and [38]. In [38], GDIs are used, while in [34] cyclostationarity and continuous wavelet transform spectrogram analysis are used for gait-based recognition. Table 3 summarizes key gait dynamics-based continuous authentication methods in terms of their performance on various data sets. In this table, the verification rate (VR), false nonmatch rate (FNMR), and false match rate (FMR) are noted in two of the studies.

Behavior-based profiling

Behavior profiling techniques verify the user's identity based on the applications and services they use. The research into mobile behavior profiling started in late 1990s, focusing mainly on developing intrusion detection systems (IDSs) to detect telephony service fraud by monitoring user calling and migration behavior [40]–[42]. In these systems, user profiles are created by monitoring user activities for a period of time and are compared against the current activity profiles of the user. If a significant deviation is observed, a possible intrusion is detected.

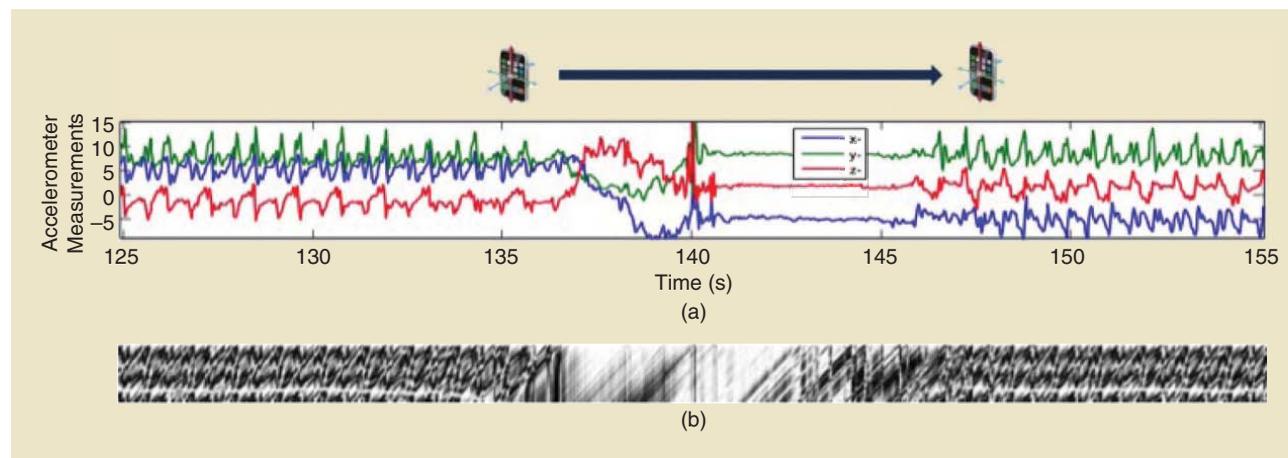


FIGURE 8. (a) Data measurements from a three-axis accelerometer embedded in a mobile phone carried by a walking user. (b) The corresponding GDI.

Table 3. Key gait-based continuous authentication methods for mobile devices.

Study	Number of Users	Feature	Classifier	Performance (%)
Mantjarvi et al. [32]	36	Raw data	Correlation coefficients	EER: 7
Thang et al. [36]	11	FFT	SVM	Accuracy: 92.7
Muaaz et al. [37]	51	Raw data	SVM	EER: 22.49–33.30
Nickel et al. [35]	48	Raw data	HMM	FNMR: 10.42 @ FMR10.29
Zhong et al. [38]	51	GDI	Nearest neighbor	EER: 3.88–7.22
Juefei-Xu et al. [34]	36	Wavelets	SVM	61.1–99.4 VR @ 0.1 FAR

Recently, a number of different techniques have been developed in the literature that focus on the use of such methods for continuous authentication [43]–[45]. In these methods, application-level as well as application-specific features such as cell ID, date, time and number of caller, duration of call, application time, and name and time of application usage are used to continuously monitor the user identity. For instance, EERs of 5.4%, 2.2%, and 13.5% have been reported in [43] for telephony, text messaging, and general application usage, respectively, on the MIT Reality data set [46]. Historical application usage data have also been utilized to verify mobile users in a continuous manner. [44] developed a technique based on historical usage data using a combination of a rule-based classifier, a dynamic profiling technique, and a smoothing function. They reported an EER of 9.8%. Recently, a behavior profiling method that focuses on what, where, when, and how the mobile devices were used was developed in [45]. A privacy-preserving implicit authentication system based on behavior profiling has also been proposed in [47]. In [48], a data-driven approach was proposed for continuous authentication based on incremental training. The researchers argued that a few weeks of data may not be sufficient for training but that training must be set automatically on a per-use basis.

Furthermore, a behavior profiling method based on application usage, Bluetooth sightings, and Wi-Fi access point sightings was recently presented in [49]. Discriminative features from these modalities were extracted, and a categorical nearest-neighbor classifier was used to produce matching scores. The authors reported average identification rates of 80%, 77%, 93%, and 85% when using application, Bluetooth, Wi-Fi, and the combination of these three types of behavioral

features, respectively. Table 4 summarizes the results for all the behavior-based profiling methods discussed above.

Other approaches

Keystroke dynamics is another behavioral biometric that is widely used to continuously authenticate mobile-device users. In keystroke dynamics, users are identified based on their typing patterns. In this method, two types of features, interkey time (the time between two successive key presses) and hold time (the time between pressing and releasing a single key), are commonly used. In particular, [50] proposed the use of keystroke dynamics based on the way users type graphical-based passwords to authenticate the mobile-device users. Some of the other keystroke dynamics-based methods can be found in [51]–[53].

Mobile-device movement and the ambient noise measured by smartphone microphones were used in [54] to implicitly authenticate mobile-device users. Based on the data captured from nine subjects, the authors reported recognition accuracy of 88.3%, 47.8%, and 90.1% for movement, audio, and the combination of these two features, respectively. Furthermore, [55] studied the feasibility of voice biometrics on mobile devices. It was shown that a mobile user's identity could be verified by his or her voice with an EER of 7.77%. In [56], linguistic profiling was used to authenticate users based on their writing vocabulary and style of short-message-service message. Experimental results based on 30 participants showed that linguistic profiling can be successfully used to authenticate users, with low error rates.

Several studies have used contextual information to enhance the performance of continuous authentication. For example, [57] investigates how the position in which the smartphone is held affects user authentication. Another

In keystroke dynamics, users are identified based on their typing patterns.

Table 4. Key behavior profiling-based continuous authentication methods for mobile devices.

Study	Behavior	Data Set (Users)	Classifier	Performance (%)
Li et al. [43]	Application usage	MIT Reality	Neural net	EER: 13.5
Li et al. [43]	Text message	MIT Reality	Neural net	EER: 2.2
Li et al. [43]	Calls	MIT Reality	Neural net	EER: 5.4
Li et al. [44]	Historical usage data	MIT Reality	Neural net	EER: 9.8
Neal et al. [49]	Application usage, Bluetooth, and Wi-Fi	UND data set (200)	Nearest neighbor	RR: 80-93

context-aware continuous authentication method [58] proposes to use passive as well as active factors to continuously authenticate users. The authors argue that digital sensors, combined with models of people and places, can give some information about user identity. In [59], contextual application information is used to improve user authentication based on touch gestures.

Fusion of multiple modalities

Unimodal continuous authentication systems rely on a single source of information, such as touch gestures, faces, or behavior profiling. Such unimodal systems have to deal with some of the following inevitable problems [60]:

- *Noisy data*: Poor lighting on a user's face or occlusion are examples of noisy data.
- *Nonuniversality*: The continuous authentication system based on a single source of evidence may not be able to capture meaningful data from some users. For instance, gait-based systems may extract incorrect patterns for certain users due to leg injuries.
- *Intraclass variations*: These often occur when a user incorrectly interacts with the sensor.
- *Spoof attack*: Using a photograph to gain access to a user's mobile device is an example of this type of attack.

It has been observed that some of the limitations of unimodal continuous authentication systems can be addressed by deploying multimodal systems that essentially integrate the evidence presented by multiple sources of information such as touch gestures and faces. Such systems are less vulnerable to spoof attacks, as it would be difficult for an imposter to simultaneously spoof multiple biometric traits of a genuine user.

Classification in multimodal systems is done by fusing information from different modalities. The information fusion can be done at different levels, which can be broadly divided into feature-level, score-level, and rank/decision-level fusion. Several methods have been proposed in the literature that make use of multiple modalities for continuous authentication. For instance, a feature-level fusion method based on multitask multivariate low-rank representations was recently proposed in [61] for fusing touch gestures and faces for continuous authentication. A decision-level fusion method was proposed in [62] for fusing four modalities based on stylometry (text analysis), application usage patterns, web browsing behavior, and physical location of the device for continuous authentication. The analysis performed on a data set of 200 Android mobile-device users whose data were collected for a period of at least 30 days showed that the method can achieve an EER of 0.05 using a one-minute window and an EER below 0.01 using a 30-minute window. Similarly, in [63] a SenGuard system was proposed in which multiple modalities are fused at the decision level

for continuous authentication. Data from the accelerometer, touch screen, and microphone as well as location history are used to continuously monitor the user identity on a mobile device. In the authors' approach, they rely on the Jigsaw continuous sensing engine [64] to process the motion and voice data. Furthermore, their touch-based method can handle single as well as multitouch gestures.

A bimodal continuous authentication method based on face and speaker recognition was proposed in [65]. The authors' face detection and recognition approach is based on LBPs [23]. For speaker recognition, voice activity detection is first performed using a Hungarian downsampled phoneme recognizer, which is essentially the cascade of three neural networks. After voice activity detection, all valid frames are passed to the speaker-authentication component, which uses an *i*-vector extractor to obtain features that are then modeled using probabilistic linear discriminant analysis. Finally, similarity scores for face authentication and the log-likelihood scores for speaker authentication are normalized to produce probabilities and fused by taking the product of the two resulting scores.

Recently, a set of behavioral features called hand movement, orientation, and grasp (HMOG) was proposed in [66] to continuously authenticate smartphone users. HMOG is essentially based on the accelerometer, gyroscope, and magnetometer readings and captures subtle hand micromovements and orientation patterns generated when a user taps on the screen. A set of 96 HMOG features was proposed and evaluated on a data set consisting of 100 users' typing data. It was shown that one can achieve authentication EERs as low as 7.16% (walking) and 10.05% (sitting) when the HMOG features are combined with tap and keystroke features using a score-level fusion framework [66]. Table 5 summarizes the key multimodal fusion methods for continuous authentication in terms of their performance on various data sets. In this table, the half total error rate (HTER) is noted for one of the studies.

In [67], three different text-based biometric modalities—linguistic profiling, behavioral profiling, and keystroke dynamics—were fused using a score-level fusion method for continuous authentication. Since there is no multimodal data set that consists of these three text-based biometric modalities for the same individual, these modalities were combined from different data sets to create a virtual data set of 30 users. Based on this data set, the authors reported an average EER of 3.3% when linguistic profiling, behavioral profiling, and keystroke dynamics are fused.

In [67], three different text-based biometric modalities—linguistic profiling, behavioral profiling, and keystroke dynamics—were fused using a score-level fusion method for continuous authentication. Since there is no multimodal data set that consists of these three text-based biometric modalities for the same individual, these modalities were combined from different data sets to create a virtual data set of 30 users. Based on this data set, the authors reported an average EER of 3.3% when linguistic profiling, behavioral profiling, and keystroke dynamics are fused.

Summary of continuous authentication approaches

As discussed previously, several physiological and behavioral biometrics-based techniques have direct application

Some of the limitations of unimodal continuous authentication systems can be addressed by deploying multimodal systems that essentially integrate the evidence presented by multiple sources of information such as touch gestures and faces.

Table 5. Key multimodal fusion-based continuous authentication methods for mobile devices.

Study	Modalities	Number of Users	Fusion Method	Performance (%)
Zhang et al. [61]	Face, touch gestures	50	Feature level	RR: 83.75
Fridman et al. [62]	Stylometry, application usage, web browsing, GPS location	200	Decision level	EER: 5 (One minute), 1 (30 minute)
Shi et al. [63]	Accelerometer, touch screen, microphone, location history	Seven	Decision level	EER: -
McCool et al. [65]	Face, voice	152	Score level	HTER: 11.9 (male), 13.3 (female)
Sitova et al. [66]	HMOG, tap, keystroke	100	Score level	EER: 7.16 (walking), 10.05 (sitting)
Saeveane et al. [67]	Linguistic profiling, behavioral profiling, keystroke dynamics	30	Score level	EER: 3.3 (weighting) - 4.4 (sum)

within a continuous authentication framework. Several research studies have specifically focused on the applicability of these biometrics modalities for nonintrusive authentication. It is seen that physiological biometrics such as the face can provide higher authentication accuracy than behavioral biometrics such as gait or touch gestures. Further, as behavioral biometric characteristics tend to change over time and under various environmental conditions, one has to constantly update the templates to maintain the performance of these techniques. The tradeoff among computation, processing speed, and accuracy has to be considered when using these modalities for transparent authentication. For example, the face-based continuous authentication system requires one to detect, align, and recognize faces from the images or videos collected from the front-facing camera. Each of these subalgorithms can be very time consuming, making the overall matching algorithm computationally demanding and not real time. In contrast, touch gesture-based methods often do not require detection or segmentation of data. Hence, they could be more efficient in terms of processing speed. It can be concluded that there is not a single biometric modality that is ideally suited for all scenarios. However, a significant amount of prior research has shown that continuous authentication methods based on multiple biometric traits are often superior to unimodal continuous authentication systems.

Usability and security issues

The usability of transparent continuous authentication systems on mobile devices has become a major issue in research [5], [68], [69]. A balance needs to be struck between security and usability of a biometrics-based continuous authentication system. The design of usable yet secure continuous user authentication systems raises crucial questions concerning how to solve conflicts between mobile security and usability. For instance, in the continuous authentication context, a false rejection is less costly than a false acceptance. This is due to the fact that higher false acceptance rates will lower

the security level of the continuous authentication system, while a higher false rejection rate will frustrate a legitimate user, which is less dangerous than a lower security level. It was argued in [70] that to be able to build reliable, effective, and usable systems, one needs specific guidelines that take into account the specific constraints of security mechanisms. Furthermore, security systems should be built so as to be easy to learn and use by users with different backgrounds and skills. It was also argued that human factors should be incorporated into the design of continuous authentication systems, where usability is central during the whole development process.

Several works have discussed the issue of usability of continuous authentication systems. For instance, in [68] a prototype was developed using keystroke, voice, and face biometrics for continuous authentication. The prototype was evaluated using 27 participants, and the study reported that 92% of the participants considered it more secure in comparison

to the traditional methods of authentication. Similarly, [69] conducted an in-lab study of security perception of implicit authentication with 30 users based on behavioral biometrics. In their study, the researchers asked users to complete a series of tasks on a smartphone that was ostensibly protected with varying degrees of transparent authentication. They then surveyed the participants regarding their opinion about transparent authentication. They found that 73% of participants felt that implicit authentication based on behavioral biometrics was more secure than traditional methods such as PINs and passwords and 90% indicated that they would consider using a transparent authentication method on their own mobile device.

More recently, a two-part study consisting of a controlled lab experiment and a field study was conducted in [5] on implicit authentication usability and security perceptions with 37 participants. The study indicated that 91% of participants found implicit authentication to be convenient and 81% perceived the provided level of protection to be satisfactory. Furthermore, the authors found that false accepts and detection

Physiological biometrics such as the face can provide higher authentication accuracy than behavioral biometrics such as gait or touch gestures.

delay were prime security concerns for 27% and 22% of the participants, respectively. Also, 35% of the participants were annoyed by false rejects. These studies show that users are willing to consider trying mobile transparent and continuous authentication methods based on biometrics, as they see a need for alternatives to secret knowledge techniques such as passwords and PINs.

Discussions and future directions

This article presented an overview of recent advances in mobile-based continuous authentication methods that included behavioral, physiological, and multimodal biometrics-based fusion methods. We hope that the survey has helped to guide an interested reader among the extensive literature to some degree. Obviously, it could not cover all the literature on continuous authentication, so we chose to focus on a representative subset of the latest progress made in biometrics and the security community. Continuous authentication on mobile devices promises to be an active area of research, especially as more and more sensors are being added to the smartphone device and the computation power of mobile devices is increasing tremendously. There are, however, several challenges to be overcome before successfully designing a biometrics-based continuous authentication system. These challenges include the following.

- 1) The biometric data at enrollment time may have different characteristics than those presented during authentication. For example, in the case of the face biometric, the enrolled faces are usually frontal and well illuminated. However, during authentication the mobile device has to process faces that may have very poor illumination, severe pose variations, or missing facial parts. This problem where the training (enrolled) data used to learn a recognition or authentication model have a different distribution from the data on which the model is applied is often known as *domain adaptation* [71]. One such method based on faces and touch gestures for continuous authentication using domain adaptation was recently proposed in [72]. Domain adaptation and transfer learning techniques can be used to deal with the changing distribution problem in continuous authentication. More domain adaptive methods for mobile-based continuous authentication are needed.
- 2) As more and more continuous authentication systems are becoming available, businesses have started to integrate these technologies into their products. Often, continuous authentication technologies are outsourced to companies that provide authentication and identity assurance as a service, because deploying and maintaining these technologies require specialized expertise and infrastructure. This raises privacy concerns, because biometric information is disclosed to a third party. To deal with this issue, methods for securely outsourcing continuous authentication systems are needed [73].
- 3) Some of the behavioral biometrics-based continuous authentication methods discussed in this article are based on very simple features. For instance, most touch gesture-based methods make use of very simple features based on the x , y coordinates and time information. However, they usually do not make use of the dynamics present in the touch gestures. We feel that incorporating the dynamics as well as the geometry of touch gestures into a feature extraction algorithm can significantly enhance the performance of a touch based continuous authentication system. Selection of appropriate features is another important problem to be addressed in continuous authentication.
- 4) Some of the physiological as well as behavioral biometrics-based continuous authentication methods are vulnerable to spoof, mimic, statistic, or digital replay attacks [74], [75]. For example, one can spoof speaker authentication systems by using voice morphing techniques. Some efforts have been made in the literature to address these issues for continuous authentication. However, more is needed. For instance, in the case of face biometrics, making use of additional sensors for liveness detection would counter the problem of spoof attacks.
- 5) Many continuous authentication methods have been proposed in the literature that evaluate the performance of their proposed method on a variety of different data sets using different performance measures. However, there is no clear standard for evaluating the performance of different methods in the literature. Guidelines on an acceptable benchmark are needed.
- 6) As discussed in the previous section, most continuous authentication methods ignore the usability and acceptability issues. Even though a few recent works have attempted to address these issues, more is needed.
- 7) Unlike credit cards and passwords, which can be revoked and reissued when compromised, biometrics are permanently associated with a user and cannot be replaced. To prevent the theft of the biometric patterns of mobile-device users, biometric template protection schemes such as cancelable biometrics [76] should be incorporated within the continuous authentication framework.
- 8) Most mobile-based continuous authentication techniques discussed in this article have been evaluated on small and midsize data sets consisting of hundreds of samples. However, to really see the significance and impact of various continuous authentication schemes in terms of usability and security, they need to be evaluated on large-scale data sets containing thousands and millions of samples.

Authors

Vishal M. Patel (vishal.m.patel@rutgers.edu) is an assistant professor in the Department of Electrical and Computer Engineering at Rutgers University, Piscataway, New Jersey. Prior to joining Rutgers University, he was a member of the research faculty at the University of Maryland Institute for Advanced Computer Studies. His research interests include

signal processing, computer vision, and machine learning, with applications to radar imaging and biometrics. He is a recipient of the 2016 Office of Naval Research Young Investigator Award and the 2010 Oak Ridge Associated Universities postdoctoral fellowship. He is member of Eta Kappa Nu, Pi Mu Epsilon, and Phi Beta Kappa. He is a Senior Member of the IEEE.

Rama Chellappa (rama@umd.edu) is a professor of electrical and computer engineering (ECE) and an affiliate professor of computer science at the University of Maryland, College Park. He is also affiliated with the Center for Automation Research, Institute for Advanced Computer Studies (permanent member), and is the chair of the ECE Department. In 2005, he was named a Minta Martin Professor of Engineering. His current research interests are clustering; three-dimensional modeling from video; image- and video-based recognition of objects, events, and activities; dictionary-based inference; compressive sensing; domain adaptation; and hyperspectral processing. He is a Fellow of the IEEE, the Association for Computing Machines, the International Association of Pattern Recognition, the Optical Society of America, the American Association for the Advancement of Science, and the Association for the Advancement of Artificial Intelligence.

Deepak Chandra (dchandra@google.com) heads authentication in the Machine Intelligence and Research group at Google Inc., Mountain View, California. The project aims to completely redefine authentication for the digital and physical worlds. Prior to this, he was the program lead in Google's Advanced Technology and Projects organization, where he headed all product engineering and design for mobile authentication projects. He defined company-wide authentication strategy for Motorola prior to leading the efforts at Google. He has developed multiple wearable authentication products, including Motorola Skip and Digital Tattoo.

Brandon Barbelo (bbarbelo@google.com) is a product manager at Google Research and Machine Intelligence, where he works on privacy-sensitive, on-device machine learning. He was previously involved with the Google Advanced Technology and Projects organization on the Project Abacus team, where he managed efforts to develop a multimodal continuous authentication system for smartphones. Prior to his time at Google, he cofounded four companies in electronics, fintech, and private equity.

References

- [1] N. Clarke and S. Furnell, "Authentication of users on mobile telephones: A survey of attitudes and practices," *Comput. and Security*, vol. 24, no. 7, pp. 519–527, 2005.
- [2] A. Vance. (2010, Jan. 20). If your password is 123456, just make it hackme [Online]. Available: <http://www.nytimes.com>
- [3] A. J. Aviv, K. Gibson, E. Mossop, M. Blaze, and J. M. Smith, "Smudge attacks on smartphone touch screens," in *Proc. 4th USENIX Conf. Offensive Technol.*, 2010, pp. 1–7.
- [4] D. Tapellini. (2014, May 28). *Smart phone thefts rose to 3.1 million in 2013: Industry solution falls short, while legislative efforts to curb theft continue* [Online]. Available: <http://www.consumerreports.org/cro/news/2014/04/smart-phone-thefts-rose-to-3-1-million-last-year/index.htm>
- [5] H. Khan, U. Hengartner, and D. Vogel, "Usability and security perceptions of implicit authentication: Convenient, secure, sometimes annoying," in *11th Symp. Usable Privacy and Security (SOUPS 2015)*, 2015, pp. 225–239.
- [6] S. Egelman, S. Jain, R. S. Portnoff, K. Liao, S. Consolvo, and D. Wagner, "Are you ready to lock?" in *Proc. 2014 ACM SIGSAC Conf. Comput. and Commun. Security*, 2014, pp. 750–761.
- [7] M. Harbach, E. von Zezschwitz, A. Fichtner, A. D. Luca, and M. Smith, "It's a hard lock life: A field study of smartphone (un)locking behavior and risk perception," in *Symp. Usable Privacy and Security (SOUPS 2014)*, 2014, pp. 213–230.
- [8] R. P. Guidorizzi, "Security: Active authentication," *IT Professional*, vol. 15, no. 4, pp. 4–7, July/Aug. 2013.
- [9] M. Jakobsson, E. Shi, P. Golle, and R. Chow, "Implicit authentication for mobile devices," in *Proc. USENIX*, 2009, pp. 1–9.
- [10] E. Shi, Y. Niu, M. Jakobsson, and R. Chow, "Implicit authentication through learning user behavior," in *Proc. 13th Int. Conf. Inform. Security*, 2011, pp. 99–113.
- [11] N. L. Clarke, *Transparent User Authentication: Biometrics, RFID and Behavioural Profiling*. London: Springer, 2011.
- [12] D. Crouse, H. Han, D. Chandra, B. Barbelo, and A. K. Jain, "Continuous authentication of mobile user: Fusion of face image and inertial measurement unit data," in *Int. Conf. Biometrics*, 2015, pp. 135–142.
- [13] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," *IEEE Trans. Inform. Forensics and Security*, vol. 8, no. 1, pp. 136–148, 2013.
- [14] L. Li, X. Zhao, and G. Xue, "Unobservable reauthentication for smart phones," in *Proc. 20th Network and Distributed Syst. Security Symp.*, 2014, pp. 1–16.
- [15] A. Serwadda, V. Phoha, and Z. Wang, "Which verifiers work? A benchmark evaluation of touch-based authentication algorithms," in *Proc. IEEE Int. Conf. Biometrics: Theory, Applications and Systems*, Sept. 2013, pp. 1–8.
- [16] H. Zhang, V. M. Patel, M. E. Fathy, and R. Chellappa, "Touch gesture-based active user authentication using dictionaries," in *Proc. IEEE Winter Conf. Applicat. Comput. Vision*, 2015, pp. 207–214.
- [17] T. Feng, Z. Liu, K. A. Kwon, W. Shi, B. Carbanar, Y. Jiang, and N. Nguyen, "Continuous mobile authentication using touchscreen gestures," in *Proc. IEEE Conf. Technol. Homeland Security*, Nov. 2012, pp. 451–456.
- [18] M. Sherman, G. Clark, Y. Yang, S. Sugrim, A. Modig, J. Lindqvist, A. Oulasvirta, and T. Roos, "User-generated free-form gestures for authentication: Security and memorability," in *Proc. 12th Annu. Int. Conf. Mobile Syst., Applicat., and Services*, 2014, pp. 176–189.
- [19] X. Zhao, T. Feng, and W. Shi, "Continuous mobile authentication using a novel graphic touch gesture feature," in *Proc. IEEE Int. Conf. Biometrics: Theory, Applicat. and Syst.*, Sept. 2013, pp. 1–6.
- [20] X. Zhao, T. Feng, W. Shi, and I. Kakadiaris, "Mobile user authentication using statistical touch dynamics images," *IEEE Trans. Inform. Forensics and Security*, vol. 9, no. 11, pp. 1780–1789, 2014.
- [21] A. Hadid, J. Heikkilä, O. Silven, and M. Pietikainen, "Face and eye detection for person authentication in mobile phones," in *Proc. ACM/IEEE Int. Conf. Distributed Smart Cameras*, Sept. 2007, pp. 101–108.
- [22] P. A. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [23] T. Ojala, M. Pietikainen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.
- [24] S. Sarkar, V. M. Patel, and R. Chellappa, "Deep feature-based face detection on mobile devices," in *Proc. IEEE Int. Conf. Identity, Security and Behavior Anal.*, 2016.
- [25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Inform. Processing Syst.*, 2012, pp. 1097–1105.
- [26] M. E. Fathy, V. M. Patel, and R. Chellappa, "Face-based active authentication on mobile devices," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2015, pp. 1687–1691.
- [27] U. Mahbub, V. M. Patel, D. Chandra, B. Barbelo, and R. Chellappa, "Partial face detection for continuous authentication," in *Proc. IEEE Int. Conf. Image Processing*, 2016.
- [28] P. Abeni, M. Baltatu, and R. D'Alessandro, "Nis03-4: Implementing biometrics-based authentication for mobile devices," in *Proc. IEEE Global Telecommun. Conf.*, Nov. 2006, pp. 1–5.
- [29] S. Chen, A. Pande, and P. Mohapatra, "Sensor-assisted facial recognition: An enhanced biometric authentication system for smartphones," in *Proc. 12th Annu. Int. Conf. Mobile Syst., Applicat., and Services*, 2014, pp. 109–122.

- [30] N. Kumar, A. Berg, P. Belhumeur, and S. Nayar, "Describable visual attributes for face verification and image search," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 33, no. 10, pp. 1962–1977, 2011.
- [31] P. Samangouei, V. M. Patel, and R. Chellappa, "Attribute-based continuous user authentication on mobile devices," in *Proc. IEEE Int. Conf. Biometrics: Theory, Applicat. and Syst.*, 2015, pp. 1–8.
- [32] J. Mantyjarvi, M. Lindholm, E. Vildjiounaite, S. M. Makela, and H. Ailisto, "Identifying users of portable devices from gait pattern with accelerometers," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, March 2005, vol. 2, pp. ii/973–ii/976.
- [33] M. Derawi, C. Nickel, P. Bours, and C. Busch, "Unobtrusive user-authentication on mobile phones using biometric gait recognition," in *Proc. Int. Conf. Intelligent Inform. Hiding and Multimedia Signal Processing*, Oct. 2010, pp. 306–311.
- [34] F. Juefei-Xu, C. Bhagavatula, A. Jaech, U. Prasad, and M. Savvides, "Gait-ID on the move: Pace independent human identification using cell phone accelerometer dynamics," in *Proc. IEEE Int. Conf. Biometrics: Theory, Applicat. and Syst.*, Sept. 2012, pp. 8–15.
- [35] C. Nickel, C. Busch, S. Rangarajan, and M. Mobius, "Using Hidden Markov Models for accelerometer-based biometric gait recognition," in *Proc. IEEE Int. Colloq. Signal Processing and Its Applicat.*, March 2011, pp. 58–63.
- [36] H. M. Thang, V. Q. Viet, N. D. Thuc, and D. Choi, "Gait identification using accelerometer on mobile phone," in *Proc. Int. Conf. Control, Automation and Inform. Sci.*, Nov. 2012, pp. 344–348.
- [37] M. Muaaz and R. Mayrhofer, "An analysis of different approaches to gait recognition using cell phone based accelerometers," in *Proc. Int. Conf. Advances in Mobile Computing and Multimedia*, 2013, pp. 293:293–293:300.
- [38] Y. Zhong, Y. Deng, and G. Meltzner, "Pace independent mobile gait biometrics," in *Proc. IEEE Int. Conf. Biometrics Theory, Applicat. and Syst.*, Sept. 2015, pp. 1–8.
- [39] Y. Zhong and Y. Deng, "Sensor orientation invariant mobile gait biometrics," in *Proc. IEEE Int. Joint Conf. Biometrics*, Sept. 2014, pp. 1–8.
- [40] P. Gosset, "Aspect: Fraud detection concepts: Final report," Tech. Rep. AC095/VOD/W22/DS/P/18/1, Jan. 1998.
- [41] Y. Moreau, H. Verrelst, and J. Vandewalle, "Detection of mobile phone fraud using supervised neural networks: A first prototype," in *Proc. Int. Conf. Artificial Neural Networks*, 1997, pp. 1065–1070.
- [42] J. Hall, M. Barbeau, and E. Kranakis, "Anomaly-based intrusion detection using mobility profiles of public transportation users," in *Proc. IEEE Int. Conf. Wireless and Mobile Computing, Networking and Commun.*, Aug. 2005, vol. 2, pp. 17–24.
- [43] F. Li, N. Clarke, M. Papadaki, and P. Dowland, "Behaviour profiling for transparent authentication for mobile devices," in *Proc. Euro. Conf. Inform. Warfare and Security*, 2011, pp. 307–314.
- [44] F. Li, N. Clarke, M. Papadaki, and P. Dowland, "Active authentication for mobile devices utilising behaviour profiling," *Int. J. Inform. Security*, vol. 13, no. 3, pp. 229–244, June 2014.
- [45] D. Bassu, M. Cochinwala, and A. Jain, "A new mobile biometric based upon usage context," in *Proc. IEEE Int. Conf. Technol. for Homeland Security*, Nov. 2013, pp. 441–446.
- [46] N. Eagle and A. Pentland, "Reality mining: Sensing complex social systems," *Personal and Ubiquitous Computing*, vol. 10, no. 4, pp. 255–268, May 2006.
- [47] N. A. Safa, R. Safavi-Naini, and S. F. Shahandashti, "Privacy-Preserving Implicit Authentication," in *ICT Systems Security and Privacy Protection*, N. Cuppens-Bouahia, F. Cuppens, S. Jajodia, A. Abou El Kalam, and T. Sans, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 471–484.
- [48] H. G. Kayacik, M. Just, L. Baillie, D. Aspinall, and N. Micallef, "Data driven authentication: On the effectiveness of user behaviour modelling with mobile device sensors," *CoRR*, 2014 vol. abs/1410.7743.
- [49] T. Neal, D. Woodard, and A. Striegel, "Mobile device application, Bluetooth, and Wi-Fi usage data as behavioral biometric traits," in *Proc. IEEE Int. Conf. Biometrics Theory, Applicat. and Syst.*, Sept. 2015, pp. 1–6.
- [50] T. Y. Chang, C. J. Tsai, and J. H. Lin, "A graphical-based password keystroke dynamic authentication system for touch screen handheld mobile devices," *J. Syst. and Software*, vol. 85, no. 5, pp. 1157–1165, 2012.
- [51] N. Clarke and S. Furnell, "Advanced user authentication for mobile devices," *Computers and Security*, vol. 26, no. 2, pp. 109–119, 2007.
- [52] N. L. Clarke and S. M. Furnell, "Authenticating mobile phone users using keystroke analysis," *Int. J. Inform. Security*, vol. 6, no. 1, pp. 1–14, Dec. 2006.
- [53] H. Gascon, S. Uellenbeck, C. Wolf, and K. Rieck, "Continuous authentication on mobile devices by analysis of typing motion behavior," presented at *Proc. GI Conf. Sicherheit (Sicherheit, Schutz und Verlässlichkeit)*, 2014.
- [54] H. Ketabdar, M. Roshandel, and D. Skripko, "Towards implicit enhancement of security and user authentication in mobile devices based on movement and audio analysis," in *Proc. 4th Int. Conf. Advances in Comput.-Human Interactions*, 2011, pp. 188–191.
- [55] R. Woo, A. Park, and T. Hazen, "The MIT mobile device speaker verification corpus: Data collection and preliminary experiments," in *Proc. Odyssey, The Speaker and Language Recognition Workshop*, 2006, pp. 1–6.
- [56] H. Saevanee, N. L. Clarke, and S. M. Furnell, "SMS linguistic profiling authentication on mobile devices," in *Proc. 5th Int. Conf. Network and Syst. Security*, 2011, pp. 224–229.
- [57] A. Primo, V. Phoha, R. Kumar, and A. Serwadda, "Context-aware active authentication using smartphone accelerometer measurements," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognition Workshops*, June 2014, pp. 98–105.
- [58] E. Hayashi, S. Das, S. Amini, J. Hong, and I. Oakley, "CASA: Context-aware scalable authentication," in *Proc. 9th Symp. Usable Privacy and Security*, 2013, pp. 3:1–3:10.
- [59] T. Feng, J. Yang, Z. Yan, E. M. Tapia, and W. Shi, "Tips: Context-aware implicit user identification using touch screen in uncontrolled environments," in *Proc. Workshop on Mobile Comput. Syst. and Applicat.*, 2014, pp. 9:1–9:6.
- [60] A. Ross and A. K. Jain, "Multimodal biometrics: An overview," in *Proc. Euro. Signal Processing Conf.*, 2004, pp. 1221–1224.
- [61] H. Zhang, V. M. Patel, and R. Chellappa, "Robust multimodal recognition via multitask multivariate low-rank representations," in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, May 2015, vol. 1, pp. 1–8.
- [62] L. Fridman, S. Weber, R. Greenstadt, and M. Kam, "Active authentication on mobile devices via stylometry, GPS location, web browsing behavior, and application usage patterns," *IEEE Syst. J.*, 2015.
- [63] W. Shi, J. Yang, Y. Jiang, F. Yang, and Y. Xiong, "Senguard: Passive user identification on smartphones using multiple sensors," in *Proc. IEEE Int. Conf. Wireless and Mobile Comput., Networking and Commun.*, 2011, pp. 141–148.
- [64] H. Lu, J. Yang, Z. Liu, N. D. Lane, T. Choudhury, and A. T. Campbell, "The jigsaw continuous sensing engine for mobile phone applications," in *Proc. 8th ACM Conf. Embedded Networked Sensor Syst.*, 2010, pp. 71–84.
- [65] C. McCool, S. Marcel, A. Hadid, M. Pietikainen, P. Matejka, J. Cernocky, N. Poh, J. Kittler, A. Larcher, C. Levy, D. Matrouf, J. F. Bonastre, P. Tresadern, and T. Cootes, "Bi-modal person recognition on a mobile phone: Using mobile phone data," in *Proc. IEEE Int. Conf. Multimedia and Expo Workshops*, July 2012, pp. 635–640.
- [66] Z. Sitová, J. Šeděnka, Q. Yang, G. Peng, G. Zhou, P. Gasti, and K. S. Balagani, "HMGO: New behavioral biometric features for continuous authentication of smartphone users," *IEEE Trans. Inform. Forensics Security*, vol. 11, no. 5, pp. 877–892, May 2016.
- [67] H. Saevanee, N. Clarke, S. Furnell, and V. Biscione, "Text-based active authentication for mobile devices," in *ICT Systems Security and Privacy Protection*, N. Cuppens-Bouahia, F. Cuppens, S. Jajodia, A. Abou El Kalam, and T. Sans, Eds. Berlin, Heidelberg: Springer, 2014, pp. 99–112.
- [68] N. Clarke, S. Karatzouni, and S. Furnell, "Flexible and Transparent User Authentication for Mobile Devices," in *Emerging Challenges for Security, Privacy and Trust*, D. Gritzalis and J. Lopez, Eds. Berlin, Heidelberg: Springer, 2009, pp. 1–12.
- [69] H. Crawford and K. Renaud, "Understanding user perceptions of transparent authentication on a mobile device," *J. Trust Manage.*, vol. 1, no. 7, pp. 1–28 2014.
- [70] C. Braz and J.-M. Robert, "Security and usability: The case of the user authentication methods," in *Proc. 18th Conf. L'Interaction Homme-Machine*, 2006, pp. 199–203.
- [71] V. M. Patel, R. Gopalan, R. Li, and R. Chellappa, "Visual domain adaptation: A survey of recent advances," *IEEE Signal Processing Mag.*, vol. 32, no. 3, pp. 53–69, May 2015.
- [72] H. Zhang, V. M. Patel, S. Shekhar, and R. Chellappa, "Domain adaptive sparse representation-based classification," in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, vol. 1, May 2015, pp. 1–8.
- [73] J. Sedenka, S. Govindarajan, P. Gasti, and K. Balagani, "Secure outsourced biometric authentication with performance evaluation on smartphones," *IEEE Trans. Inform. Forensics and Security*, vol. 10, no. 2, pp. 384–396, Feb. 2015.
- [74] D. F. Smith, A. Wiliem, and B. C. Lovell, "Binary watermarks: A practical method to address face recognition replay attacks on consumer mobile devices," in *Proc. IEEE Int. Conf. Identity, Security and Behavior Anal.*, Mar. 2015, pp. 1–6.
- [75] D. F. Smith, A. Wiliem, and B. C. Lovell, "Face recognition on consumer devices: Reflections on replay attacks," *IEEE Trans. Inform. Forensics and Security*, vol. 10, no. 4, pp. 736–745, Apr. 2015.
- [76] V. M. Patel, N. K. Ratha, and R. Chellappa, "Cancelable biometrics: A review," *IEEE Signal Processing Mag.*, vol. 32, no. 5, pp. 54–65, Sept. 2015.



Thomas A. Baran,
Richard G. Baraniuk,
Alan V. Oppenheim,
Paolo Prandoni, and
Martin Vetterli

In higher education circles, 2012 may be known as the “year of the MOOC”; the launch of several high-profile initiatives, both for profit (Coursera, Udacity) and not for profit (edX), created an electrified feeling in the community, with massive open online courses (MOOCs) becoming the hottest new topic in academic conversation. The sudden attention was perhaps slightly forgetful of many notable attempts at distance learning that occurred before, from campus TV networks to well-organized online repositories of teaching material. The new mode of delivery, however, was ushered in by a few large-scale computer science courses, whose broad success triggered significant media attention [1].

In the debate, some were quick to predict the end of traditional brick-and-mortar universities and marveled at the eagerness with which hallowed institutions wanted to be in the game. Some hailed the free and open MOOCs as the great equalizer, capable of providing underprivileged learners around the world with access to the

Digital Object Identifier 10.1109/MSP.2016.2556004
Date of publication: 1 July 2016

highest-quality education. Of course, after a few sobering years of experience, neither extreme actually happened; in this article, we share our collective opinion on this topic.

Introduction

As teachers of various signal processing courses, we felt intrigued by the opportunity to bring our subject to this new teaching paradigm. Of course, many online resources for signal processing education exist already, including video courses developed in the 1980s by the Massachusetts Institute of Technology (MIT) Center for Advanced Engineering Studies [2], [3], various courses available on MIT's OpenCourseWare (OCW), a variety of online interactive demos, hands-on tutorials and code samples, and, of course, many collections of lecture notes and slides. Still, there was no interactive, full-fledged online course on signal processing available as the MOOC revolution was unfolding. This presented an interesting challenge and opportunity.

We approached online teaching from different backgrounds and with different perspectives, and we all started from our respective residential classes that range from mandatory undergraduate-, to master of science- and doctoral-level courses. Specifically, during the last three years, we created

three MOOCs based on the following residential courses:

- Signal Processing for Communications (COM303), a mandatory undergraduate course taken in the third year at École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, became Digital Signal Processing (DSP) on Coursera and has been offered five times since spring 2013.
- Discrete-Time Signals and Systems (ELEC301), a mandatory undergraduate course taken in the second/third year at Rice University, is the basis for the corresponding edX course offered twice since fall 2014.
- Discrete-Time Signal Processing (6.341), a first-year graduate course at MIT that is the basis for 6.341x, made available on edX and offered twice since fall 2014: once in a limited industrial beta format and once as a fully open online course.

The goal of this article is to relate the experience gained in moving “on-campus” material to the MOOC format and share the lessons learned, the influence this has had on our on-campus teaching. We hope that our MOOC experiences are valuable to the signal processing community.

EPFL's Digital Signal Processing on Coursera

Overview and goals

The course Digital Signal Processing, by Paolo Prandoni and Martin Vetterli, was first offered on the Coursera platform in February 2013. At the time of this writing, the class has completed its fifth edition. We authors are with the EPFL, and the course is based on the residential class COM303 offered by the Communication Systems Department to third-year undergraduates. For many SysCom students at EPFL, COM303 represents the first exposure to a higher-level engineering class after two years focused primarily on introductory subjects. COM303 lists calculus and linear algebra as prerequisites and recommends familiarity with probability theory as well. The class is based on the freely available textbook, *Signal Processing for Communications* [4], that we have written.

COM303 is a standard undergraduate-level DSP class with a slight emphasis on telecommunication systems. The syllabus starts off in the discrete-time domain and uses vector spaces and linear algebra as the framework to introduce signals and signal transforms; as subsequent topics are introduced, the goal is to strike a balance between solid mathematical foundations and practical applications. When adapting COM303 to the online medium, we decided to closely mirror the residential class. We did this for two reasons: primarily, we wanted to produce a package that, although aimed at the general public, would retain its focus on theoretical foundations rather than deliver yet another hands-on approach to applied DSP, for which countless tutorials are available on the Internet. Additionally, we wanted to experiment with the concept of the “flipped classroom” and be able to minimize standard lecturing to the advantage of more question-and-answer (Q&A) interaction with on-campus students.

Course organization

Outline

COM303 is composed of 17 lecture days that occur during nine weeks, and, as shown in Table 1, it is structured around nine thematic modules; each module is split into a varying number of small units (the actual videos) with the intent of balancing the conflicting requirements of a fine-grained subdivision of the material with the “narrative” needed to provide reasonably self-contained mini-lectures. Each lecture day provides students with the following:

- three video units (with associated slides; we should mention that we found it very difficult to produce videos lasting ten minutes or fewer, as per the recommended best practices, and average video length is 17 minutes)

Table 1. The syllabus for the EPFL course.

Module	Number of Units
Introduction	1
Discrete-Time Signals	3
Hilbert Space	3
Fourier Analysis	10
Linear Filters	12
Interpolation and Sampling	6
Stochastic SP	3
Image Processing	6
Digital Communication Systems	6

- additional material in the form of a numerical example or a mini-lecture on signal processing applications (see the section “Course Evolution”)
- an automatically graded homework set; the passing grade for the class is determined from the cumulative homework score.

Overall, the course delivers 14 hours of video lectures (using approximately 1,300 slides) and 126 graded quizzes.

Style and format

It is said that Pythagoras would impart his lectures hiding behind a curtain, so that his students would concentrate solely on his words. His teaching style was called *acousmatic*, a word indicating an intelligible sound whose source remains unseen. In the same spirit, we decided to produce streamlined video lectures by pairing a slideshow with a simple voiceover; dynamic annotations drawn by hand are used to underscore key passages and elucidate derivations. Production-wise, this choice also enabled us to record and edit the audio in an efficient way before “filming” the video annotations. Great effort has been placed into the design of a large number of illustrations. To achieve a “coherent visual grammar” in the illustrations, we designed a LaTeX package called DSPTricks to efficiently draw one-dimensional (1-D) and two-dimensional DSP figures in PostScript [5]. The package is a high-level graphics toolbox that programmatically produces parametric images from within the LaTeX document and allows users to easily repurpose and modify their illustrations; we refer readers to the supplementary material that appears in IEEE *Xplore*. Finally, to preserve the “human component” of the class, brief introductions and closing remarks are added to each module in the form of short videos, in which the instructor appears in person.

Homework and grading

Homework and exercises present multiple challenges to teaching signal processing online because theoretical proofs and free-form derivations remain beyond the scope of automated graders (although some progress is underway [S2], [6]); whereas peer review may work effectively in less technical subjects, we believe that in our case the pool of students with sufficient mastery of the subject would be too small to ensure the necessary critical mass. As a consequence,

we only designed graded homework with either multiple-choice or numerical answers. This somewhat limits the palette of questions that can be effectively formulated but has the advantage of providing unambiguous results, which is important given that the passing criterion for the class is based exclusively on homework grades. To complement the homework, we provide solved problem sets, in which we tackle more articulated questions whose answers require derivations and proofs.

Course evolution

The first edition of the online class was offered in February 2013, at the height of MOOC hype. Although enrollment was definitely massive, so too was the dropout rate. The initial version of the class was produced under tight deadlines and required substantial effort and overtime. Successive editions (the fifth run ended in December 2015) have refined the original material in several respects, based on accumulated attendance data and feedback from students. We are currently in the process of reworking the material to reformat the class as a potentially self-paced course; a major part of the operation involves modifying the structure of the videos to fit in with the growing trend of bite-sized lectures.

Numerical examples

One of the most powerful features of DSP formalism is its independence from any specific programming language; this flexibility, however, also proved to be somewhat of a liability for online teaching. We knew that we wanted to provide working code with which students could play, but we also tried not to endorse one programming language specifically. To remain language agnostic, we realized that we could not assign programming homework because no realistic autograder could be put in place. Consequently, we initially decided to simply complement our lectures with a number of worked-out numerical examples, ranging from simple illustrations of signal processing algorithms to more ambitious mini-lectures with a clear focus on implementation. Originally, in the interest of expediency, we used MATLAB and encouraged students to translate the examples into their language of choice. This was aided by Mathworks’ offer to provide a complimentary student license to all enrolled students. Starting with the fourth edition, we transitioned to a fully open-source solution and migrated all of our examples to Python by way of IPython notebooks [7], [8] (Figure 1). The notebooks allow us to write examples that can be either read as a worked-out exercise or downloaded, modified, and run locally by the students. This has been very well received, so much so that in the last edition we introduced a graded numerical homework in Python, where students have to code missing blocks in a fully functional MP3 encoder.

Personnel

The first edition of the class obviously required the greatest effort. We were fortunate to be able to rely on a great team of graduate students to develop exercises and troubleshoot both

Signal of the Day Series (EPFL)

As we were preparing the third edition of the École Polytechnique Fédérale de Lausanne Digital Signal Processing (DSP) massive open online course, we thought it would be cute to start each lecture with a signal of general interest, tell its history, and show some notable examples of relevant processing. We went the extra mile to try to find “original” signals, either for their historical importance or with respect to their relevance to concepts taught in the class. The challenge was then to transform a cute idea into an attractive and engaging three- to five-minute video (or an IPython Notebook) that we call the *Signal of the Day* series.

There was no shortage of ideas, and colleagues from the lab suggested examples ranging from the mundane to the downright bizarre; after all, signals are everywhere. Our goal, in all of this, was to reach a broader community than “just” those involved in signal processing; we wanted a collection of signals from different scientific communities and those that involved diverse processing challenges. We are convinced, after all, that a lot of people in science do signal processing without realizing it; reaching out to these communities is both interesting and fun.

As an example, while talking to a German environmental scientist, we discovered that in 1821, Johann Wolfgang von Goethe, the famous German writer, had started taking daily temperature measurements in his hometown of Jena. This practice has been carried on to the present day by the Jena weather station (with only a few exceptions during World War II), and so Goethe’s time series is probably one of the oldest “live” discrete-time records in existence [22]. We turned this into our first “signal of the day,” applying a simple moving average filter to demonstrate that, despite claims by climate change skeptics, temperature is indeed rising (Figure S1).

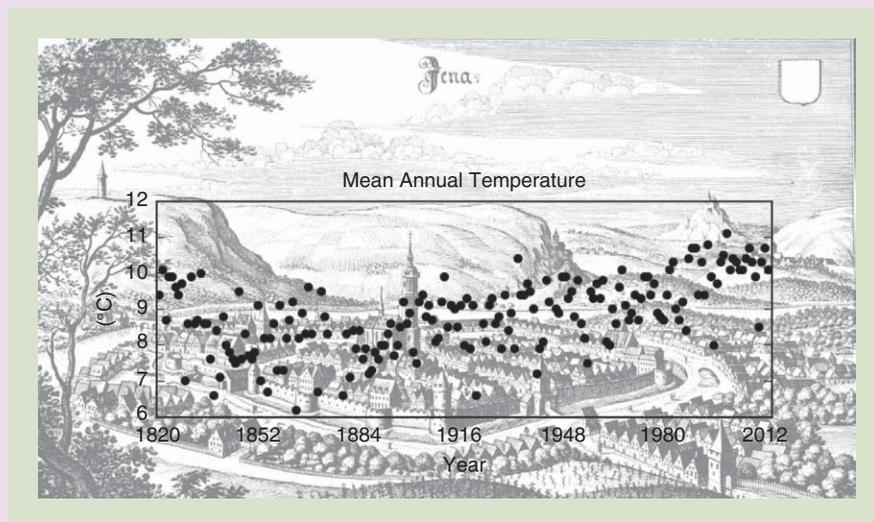


FIGURE S1. The Goethe time series.

Table S1. Signals of the day.

von Goethe’s Temperature Measurements
 The Tristan Chord
 Lehman Brothers’ stock value ca. 2008
 Sputnik: The first man-made signal from outer space
 I Feel Fine: The Beatles and guitar distortion
 Can one hear the shape of a room? [10]
 Moiré patterns
 Camera resolution and space exploration
 Exoplanet hunting
 Safecast: Crowdsourced radioactivity measurements after Fukushima [9]

Currently, we have ten signals of the day, with a few more under construction (see Table S1). A fun one is certainly the one explaining why one should never wear a striped shirt when on television, a playful way to introduce aliasing and its effects. It was also interesting to discover that other communities have developed similar ideas (see, for instance, the “astronomy picture of the day” [23]); as signal processing moves more and more to the online medium, an ever-growing library of notable DSP examples, with contributions from the entire DSP community, would be a fantastic project.

Engagement

For a more dynamic perspective, the engagement data over time for the fourth edition of the class are shown in Figure 2. The curves are very similar in shape for all editions and allow us to draw the following quantitative analysis:

- Attendance measured as passive visits to the website displays a steep decay in the first weeks of class and levels off after

week four. This roughly corresponds to the end of the Fourier analysis module, i.e., those who survive Fourier seem to keep their interest alive throughout the rest of the course.

- The percentage of students who submit homework seems to remain constant, i.e., from the start the number of students who attempt to obtain a certificate represents a small percentage of the total number of participants.

Table 2. Attendance figures on Coursera for the EPFL DSP class, with “Yield” representing passed/registered in percentage.

Edition	Date	Registered	Active	Passed	Yield
First	Feb. 2013	48,000	24,000	1,500	3.1
Second	Oct. 2013	35,000	20,000	1,000	2.9
Third	May 2014	19,000	10,000	280	1.5
Fourth	Jan. 2015	25,000	17,000	450	1.8
Fifth	Oct. 2015	16,000	7,500	360	2.2

- A peak of activity at the end of the class (April 2015) occurs, but only in terms of visits. Apparently, in online classes as in real classes, in the end students try to cram (and then give up).
- Interestingly enough, visits do not taper off to zero after the end of the class (April 2015 in Figure 2) but remain at the same level until the next offering of the course (homework submissions and forum participation obviously do stop). This strongly indicates that a MOOC model based on self-paced learning certainly has its place alongside monitored editions of the class.

Demographics

(Note: Although the following data refer to the fourth edition of the class, no significant differences have been remarked in the latest edition.) The student population showed a pronounced gender imbalance, with an 87% male component; a majority (42%) was in the 25–34 age range; and approximately half of the attendees were full-time employees and 37% full-time students. Overall, 31% held an M.S. degree and 34% a bachelor's degree, which suggests many take online classes of this kind more as a sort of refresher. The geographical distribution of students sees the United States at the top with 20%, followed by India (17%) and China (8%). As a whole, Asia leads with 38% of enrollments, followed by Europe (26%) and North America (25%).

Impact on the EPFL residential course

Initially, we thought about using MOOC recorded material to flip the classroom on campus. We attempted this during the first run of the online class, which coincided with a scheduled offering of the residential course. Students were instructed to watch the video lectures at home and prepare for Q&A sessions during the nominal lecture hours, but the experiment was unsuccessful. Students expressed an unmitigated dislike toward the absence of standard lecture time and considered their learning experience to be incomplete. We therefore decided not to repeat the format, and now we simply recommend that students enroll in the online classes and use the material to review and catch up on the standard lectures.

Feedback

Overall feedback from the online students was decidedly positive for every edition. The consensus held that the class was hard (“harder than I anticipated” was perhaps the most

common commentary in the exit questionnaire), but the rigor was almost unanimously appreciated by the students who made it to the end of the class. In general, the more senior and more educated students tended to ask for more material and a longer, in-depth class. Younger participants advocated splitting the class into shorter independent units. We certainly cannot claim to have pleased everyone, and we had our share of constructive and nonconstructive criticism. But the real privilege, as teachers, is the wealth and diversity of direct feedback that a MOOC provides. Considering that approximately 70 students per year attend the residential class, we can now sift through a century's worth of class evaluations!

Rice University's Discrete-Time Signals and Systems on edX

Overview and goals

Discrete-Time Signals and Systems (301x) is a rigorous mathematical introduction to signal processing modeled on 50% of the Rice University course ELE301, Signal and Systems, a core undergraduate class taken by all electrical and computer engineering (ECE) majors, typically in the junior year. Rather than following the classical approach to teaching discrete-time signals and systems as discretized versions of continuous-time signals and circuits, the course approaches discrete-time signals and systems from first principles. The key overarching theme is the importance of linear algebraic concepts in signal processing: vector spaces, signals as vectors, linear systems as matrices/operators, linear time-invariant (LTI) systems as Toeplitz/circulant matrices, and the Fourier transform from the eigendecomposition of these LTI matrices. The course, which also covers the z transform and filter analysis and design, teaches students to analyze discrete-time signals and systems in both the time and frequency domains. Students continuously apply these concepts in interactive MATLAB programming exercises. The course has been taught twice on edX: the first edition in spring 2014 as a ten-week course [19] and the second edition in spring 2015 as two five-week mini-courses [14], [20].

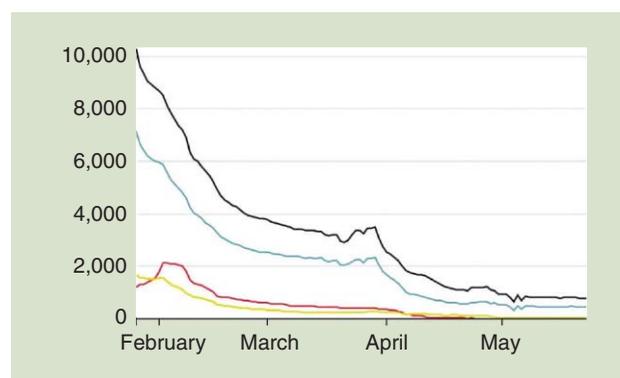


FIGURE 2. Engagement data for the fourth edition (19 January 2015–6 April 2015) of the EPFL course: visits to the course page (black), video views (blue), homework submission (red), and forum browsing (yellow).

Table 3. Topics by week in the first edition of Rice University's course, ELEC301x.

Weeks	Topics
0	Preclass activities (optional)
1	Introduction
2	Signals are vectors
3	Linear systems
4	Convolution
5	Discrete fourier transform
6	Discrete-time fourier transform
7	z transform
8	Analysis and design of filters
9	Exam

Course organization

Outline

The class incorporates numerous learning elements to engage students, stressing the balance between rigorous mathematical theory and hands-on practical applications. The course flow of the first edition is detailed in Table 3. The second edition was split into two mini-courses, one covering time-domain tools and one frequency-domain tools (with the split occurring at week 5 in Table 3).

Both course editions include an optional one-week pre-course refresher on the key prerequisites in mathematics (complex arithmetic and linear algebra) and programming (MATLAB). A collection of reference material was made available in the Rice University-based open access education platform OpenStax CNX [21]. See Table 4 for an overview of the key course elements.

Style and format

Our perusal of the cognitive science literature indicated that a “talking head” video lecture did not lead to improved learning outcomes in an online course, and so we produced lecture videos consisting of the voice of the instructor as he manipulated the slides and MATLAB windows on a tablet. See Figure 3 for sample screenshots from the course. To personify the course, the instructor appeared in a light-hearted video introducing each week's concepts. To broaden student experience and supplement the course, we produced a range of *Office Hours* videos conducted by Rice University graduate student Raajen Patel. Video production support was provided by Rice Online, a major MOOC initiative of Rice University.

Homework and grading

On the theory and analysis side, given the current limitations of the edX platform, we assess students primarily using multiple-choice questions. In the second edition of the course, each weekly homework contains one open-form response question whose response is input via MathJax and peer graded by three other students. A model solution and grading rubric are made available after each homework is due. And as discussed next, each homework also includes numerical problems in MATLAB that are assessed algorithmically via the edX platform. The final student grade combines performance on the weekly homework, case studies (recall Table 4), and the final exam. A score of 60% is required to pass the course.

Course evolution

After research by the Rice Online team suggested that MOOCs are more successful when they are shorter rather than

Table 4. Course elements in Rice ELEC301x in its first edition. The second edition split the course into two minicourses (breaking at week 5) and integrated the case studies into the weekly homework.

Precourse math refresher	Students can self-review the required mathematical skills with practice exercises and tutorials before the start of the formal course.
Preclass MATLAB tutorial	Introductory tutorials on the MATLAB programming language enable students to get up to speed.
Introduction videos	Each week kicks off with a light-hearted overview of the week's material featuring Richard Baraniuk, Mr. Lan, and BIBO, the bear.
Lecture videos	Each week features several hours of lecture videos recorded by Baraniuk specifically for an online format. Lectures were chunked into 5- to 20-minute segments.
Quick questions	Conceptual knowledge-check questions follow each segment of lecture video to test students' understanding and maintain engagement.
Supplemental resources	Links to additional learning content, exciting related applications, and additional information add depth to the learning experience.
<i>Office Hours</i> videos	Videos of a TA working out homework-type problems prepare students for the assignments and encourage appropriate problem-solving techniques.
Homework	Each week, a rigorous problem set challenges students to apply what they have learned, graded via multiple choice and peer review. For numerical problems, students program in the MATLAB language using an integrated development environment built into the edX platform.
Discussion forum	An active discussion forum enables students to ask and answer questions and receive feedback from course staff.
MATLAB case studies	Biweekly programming case studies enable students to apply concepts learned to practical programming exercises designed to show how signal processing is used.
Final exam	A traditional final exam tests students' comprehensive course knowledge.
Final case study	Serving as a final project, the final case study expands on the previous case studies as students gain programming proficiency.

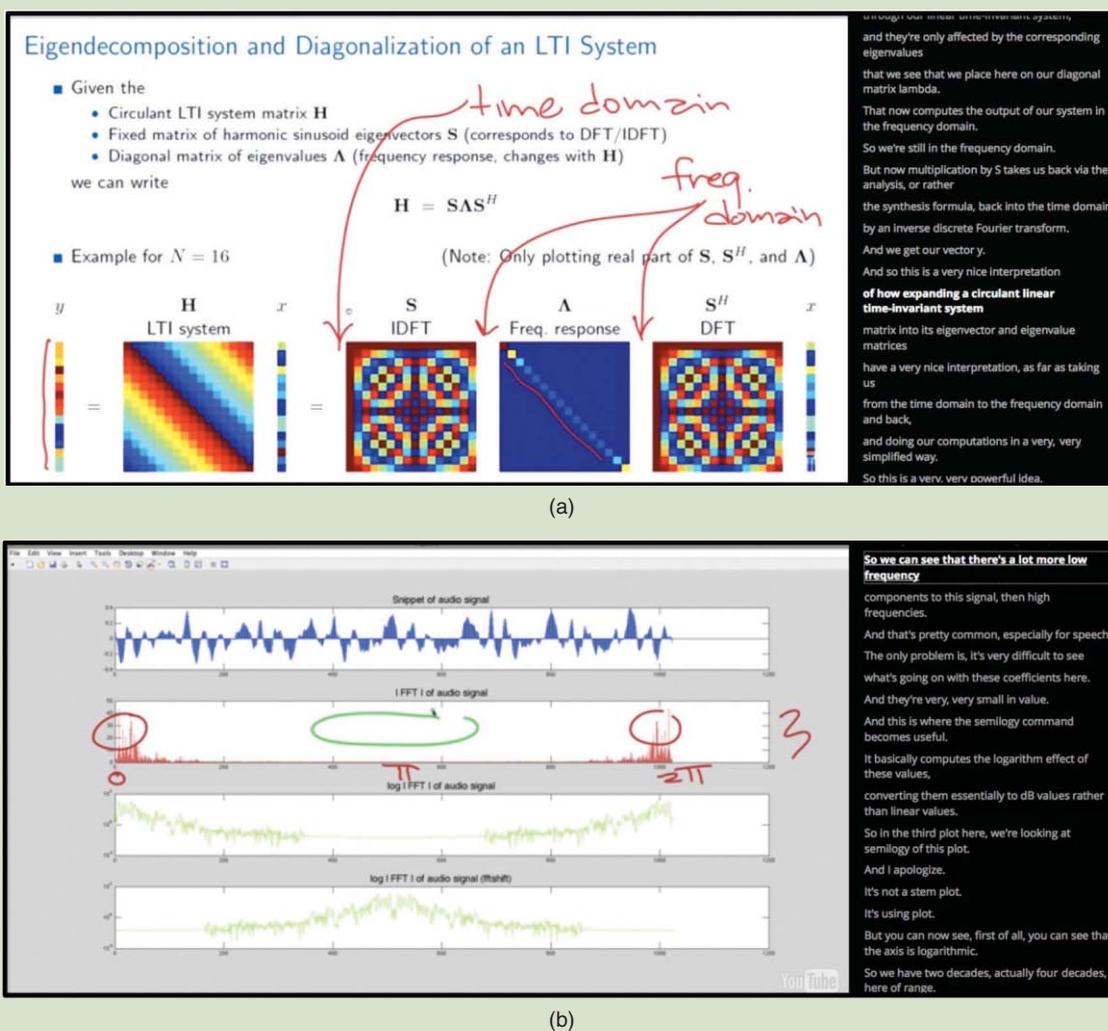


FIGURE 3. Sample screenshots from Rice University ELEC301x lecture videos showing (a) a content slide and (b) a MATLAB demonstration.

longer, the Rice administration recommended that we split the first-edition MOOC into two mini-MOOCs, each lasting five weeks. The split was accomplished by breaking the original course into one mini-course covering time-domain concepts and one mini-course covering frequency-domain concepts.

Numerical examples

As discussed previously, we provide a variety of opportunities for students to explore the numerical/application side of signal processing through 1) MATLAB examples interspersed in the lecture videos, 2) *Office Hour* videos, 3) case studies, and 4) automatically graded MATLAB problems in homework.

Personnel

Realizing the MOOC took a village, instructor Richard Baraniuk was assisted by Rice DSP research engineer Heather Seeba, DSP consultant Matthew Moravec, graduate students Raajen Patel (*Office Hours*) and Eva Dyer (case studies),

undergraduate student Tan Nguyen (MATLAB homework), and six undergraduates (forum moderation).

Unique elements

Week 0

Because students come to the class with wildly varying backgrounds, both editions of the course include an optional one-week precourse refresher on the key prerequisites in mathematics (primarily complex arithmetic and linear algebra) and programming (MATLAB). Students can test their preparedness with a pretest (for zero credit).

Integrated MATLAB

ELE301x was one of the first MOOCs to exploit a collaboration with MathWorks that made MATLAB freely available to students. Through the edX platform, students can access, within their web browser, a server running MATLAB and even submit their code for autograding.

Community Teaching Assistants (Rice University)

Conventional residential courses are typically staffed by one or more teaching assistants, who grade and sometimes create homework and exams, hold office hours, and facilitate laboratories. The concept is typically also ported over to online courses such as massive open online courses (MOOCs). The Rice ELEC301x discussion forum quickly became vibrant in both editions of the course. Moreover, certain students began taking a leadership role when answering forum questions and offering advice. One of these students, John Coppens (see <https://www.jcoppens.com>), a practicing engineer with a wealth of real-world experience, was so active in the first edition of the class that we elevated him onto the course team as a community teaching assistant for the second edition.

This is an intriguing exemplar of how the “openness” of an MOOC can lead to emergent behaviors that are unseen in residential courses. Moreover, this is a preliminary indication that the dream of MOOC scalability

could in fact be realizable. With so many of the ELEC301x students having advanced degrees (see the section “The Class by the Numbers”), we plan to continue to encourage and reward such positive contributions in the future.

John Coppens’s reflections on ELEC301x provide a number of insights into why he took such initiative: “I was very interested in the subject, and, over the years, have been studying and implementing small projects for myself. The course coincided with the first semester of the year, when I normally have a little more time to spare, and the general feel of the course through its videos and discussion forum was ‘open,’ ‘stimulating,’ and ‘inviting.’ Moreover, reactions to problems posted to the forum were prompt and helpful. Minor issues with the tasks and exercises actually forced me to do more investigation than was called for in the course, which was a great learning experience. Perhaps MOOC instructors should leave some such issues on purpose.”

Case studies

In addition to the usual analysis and calculation assessment, Rice graduate student Eva Dyer developed a suite of case studies that challenge students on the core concepts of the class through a series of real-world application programming exercises. These studies enable students to explore first hand how signal processing is used in a wide range of real applications. Case study topics include audio synthesis using sinusoids, predicting financial time series using moving average filters, audio synthesis of a clarinet using attack-decay-sustain-release curves, filter design for Karplus–Strong string synthesis, and neural spike sorting. In the second edition of the course, case studies were integrated into the weekly homework.

Office Hours videos

Rice graduate student Raajen Patel prepared videos of worked examples that give students more insight into the material and encourage strong problem-solving skills.

Post-MOOC community

Because interest in both editions of the course continued even after they were closed and archived, we worked to engage students in an ongoing, post-MOOC community. Many students were interested in staying in touch with both course staff and other students. One particularly useful contribution from 200 members of the community was to work out step-by-step solutions to a number of signal processing problems. These solutions were then used as test data for a Rice project on mathematical language processing [S2] that aims to automatically grade mathematical calculations and provide appropriate

feedback to students. For two additional unique elements, see “Community Teaching Assistants (Rice University)” and “MOOCs as an Experimental Platform (Rice University),” respectively.

The class by the numbers

We ran an interesting experiment purely by accident. For the second edition, very close to when standard practice would dictate that we announce the course (several months in advance), we were strongly recommended to split the ten-week course into two five-week courses. Accomplishing the split took several months, and so the course was announced late (approximately one month in advance). This reduced the number of registrants significantly (see Table 5). However, the proportion of students that actually completed at least the first part of the course was more than twice that in the first edition. This is evidence for the hypothesis that registration numbers are not very informative for MOOCs, because many potential students will register after reading the announcement of a course without ever truly intending to complete the work required to finish it.

Retention and engagement

The yield of enrolled to passed students was above the average for edX courses at the times our two editions were offered; see Table 5.

Demographics

The median age of residential Rice students taking ELEC301 is 20, but the median age for students taking the two editions of the course was 27. Interestingly, 30% of registrants had a high school diploma or less, 39% had a

MOOCs as an Experimental Platform (Rice University)

To help support the burgeoning learning analytics and cognitive science research program at Rice (see <http://openstaxtutor.org>), we used ELEC301x to conduct a range of experiments.

Learning analytics experiment

A study in the final week of the second edition of ELEC301x assessed whether students believe in or agree with learning analytics data that were presented to them in a dashboard. (Learning analytics involves measuring, collecting, and analyzing data about learners to understand and optimize learning.) More specifically, we aimed to determine whether students would use the analytics to guide future learning activities, the analytics were “better” than the students’ own metacognitive judgments, and our sparse factor analysis for learning and content analytics [S1] was accurate for all students. A research paper on our findings is in preparation.

Mathematics language processing experiment

The weekly open-form response question in the second edition of ELEC301x collected valuable test data for a Rice project on mathematical language processing (MLP) [S2] that aims to autograde mathematical computations and provide appropriate feedback to students. These data were augmented with additional step-by-step solutions from the 200 members of the post-MOOC community.

MLP leverages solution data from a large number of learners to evaluate the correctness of their solutions, assign partial-credit scores, and provide feedback to each learner on the likely locations of any errors. MLP takes inspiration from the success of natural language processing for text data and comprises three main steps.

- 1) Convert each solution to an open response mathematical question into a series of numerical features.
- 2) Cluster the features from several solutions to uncover the structures of correct, partially correct, and incorrect solutions.

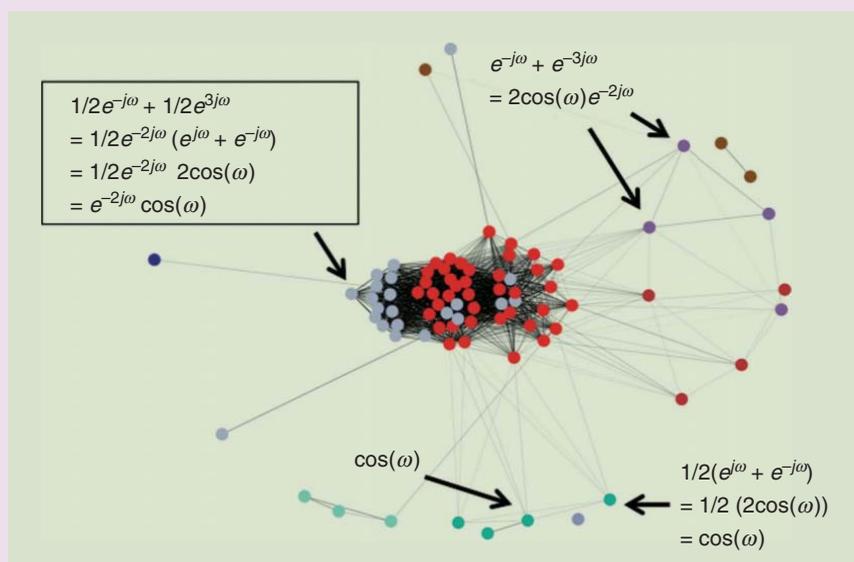


FIGURE S2. An illustration of the clusters obtained by MLP to 100 students’ solutions to a signal processing problem. Each node corresponds to a solution. Nodes with the same color correspond to solutions that are estimated to be in the same cluster. The thickness of the edge between two solutions is proportional to their similarity score. The boxed solution is correct; all others are in varying degrees of (in)correctness.

- 3) Autograde the remaining (potentially large number of) solutions based on their assigned cluster and one instructor-provided grade per cluster.

As a bonus, we can track the cluster assignment of each step of a multistep solution and determine when it departs from a cluster of correct solutions, which enables us to indicate to learners the likely locations of errors. Figure S2 illustrates the clusters of (correct and incorrect) solutions to the following signal processing problem.

Question: A discrete-time linear time-invariant system has the impulse response shown in the figure (omitted). Calculate $H(e^{j\omega})$, the discrete-time Fourier transform of the impulse response $h[n]$. Simplify your answer as much as possible until it has no summations.

References

- [S1] A. S. Lan, A. E. Waters, C. Studer, and R. G. Baraniuk, “Sparse factor analysis for learning and content analytics,” *J. Mach. Learning Res.*, vol. 15, June 2014.
- [S2] A. S. Lan, D. Vats, A. E. Waters, and R. G. Baraniuk. (2015, Jan. 18). Mathematical language processing: Automatic grading and feedback for open response mathematical questions. *ACM Learning at Scale*. [Online]. Available: <http://arxiv.org/abs/1501.04346>

college degree, and 28% had an advanced degree. As with the EPFL and MIT courses, the student population was predominantly male (83%).

The geographic distribution of the first edition of the MOOC was India (25%), United States (20%), United Kingdom (3%), Germany (3%), and China (3%), followed by 168

Table 5. Enrollments in Rice ELEC301x on edX in spring 2014.

Edition	Enrollment	Active	Passed	Yield (%)
First, spring 2014	22,819	1,145	583	2.6
Second, spring 2015, part 1	5,522		302	5.5
Second, spring 2015, part 2	4,376		161	3.7

In edX, an active student is one who submitted an assessment during the second week of the course and received a nonzero score.

other countries. In the second edition, the United States and India switched top places.

Impact on the Rice residential course

Rice ELEC301 is taught to approximately 40–60 students once per year. Our desire to make signal processing as accessible as possible in the MOOC led to a redesign of how discrete-time concepts are taught; in particular, we increased focus around linear algebra. This new approach was well received by the

Rice students in fall 2015. Moreover, on-campus students took advantage of the recorded lectures as a supplemental resource and were disappointed that the lectures were unavailable for the continuous-time portion of the class.

Feedback

We conducted a survey at the conclusion of both editions of the course using Google Forms and Qualtrics. A sampling of the results from 628 responses from the first edition is given in Figure 4. Students appreciated the vector-space approach to signals and systems, saying that it made the key concepts more accessible. Students also appreciated the optional week 0 refresher on the required mathematics and MATLAB.

MIT’s Discrete-Time Signal Processing on edX

Overview and goals

The MIT Discrete-Time Signal Processing MOOC 6.341x on edX.org [15], coauthored by Alan Oppenheim and Tom Baran, is an outgrowth of and very strongly parallels the

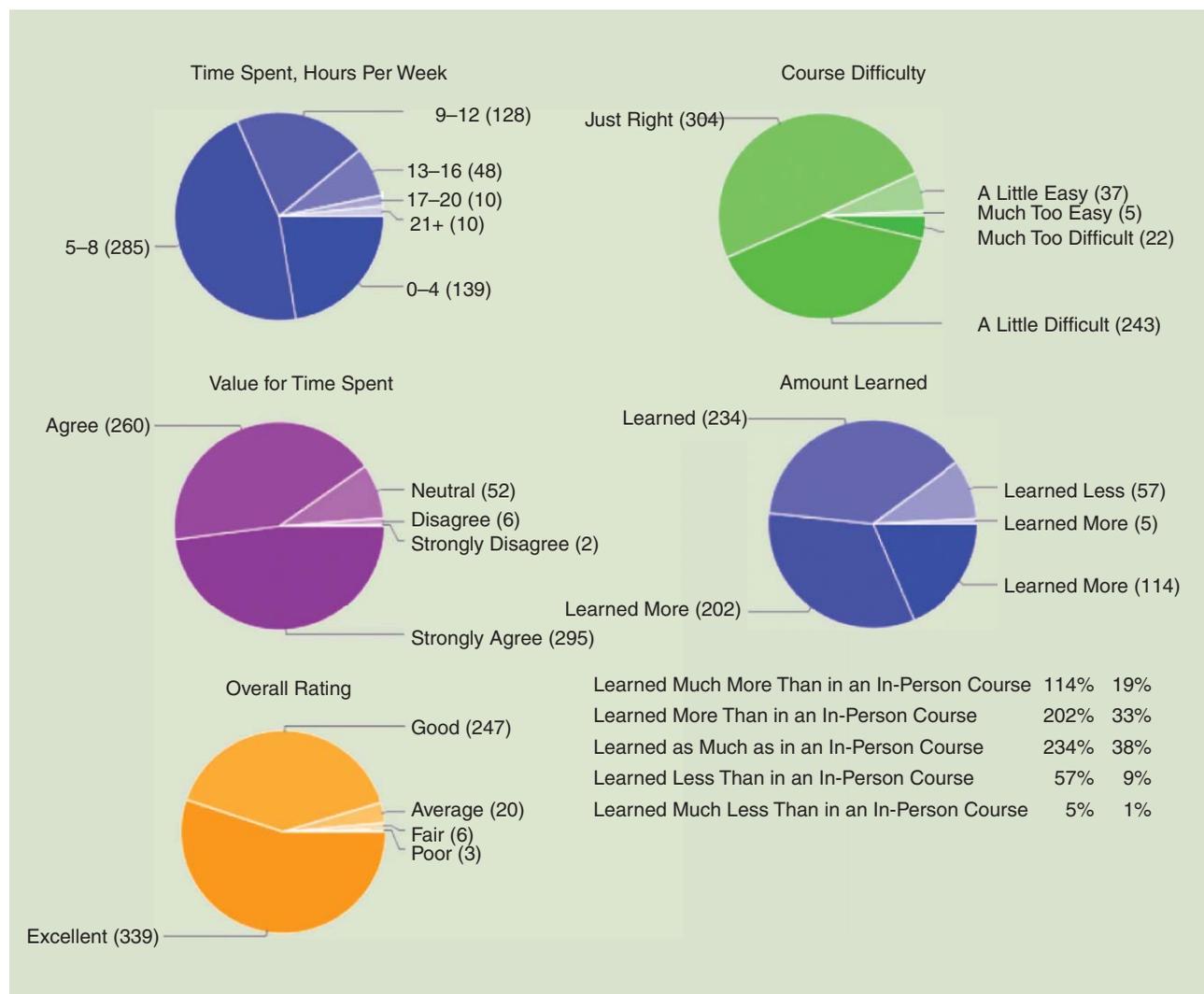


FIGURE 4. A postcourse survey of Rice University 628 ELEC301x students after the first edition.

MIT residential introductory graduate subject 6.341, the textbook for which is Discrete-Time Signal Processing [17]. The residential course carries as a prerequisite the undergraduate subject 6.011, the textbook for which is Signals, Systems, and Inference [18], which itself has as a prerequisite of the basic background of Signals and Systems.

The residential course 6.341 has a lineage and evolution at MIT that go back to the 1970s, and a key goal in creating 6.341x has been to extend that lineage and evolution forward into the online realm. It is commonly expressed that the MOOC and online teaching and learning landscape is currently “the wild west.” The territory is rapidly evolving and currently in a highly experimental stage. With this in mind, our overall goals with 6.341x included experimenting with online teaching and the edX platform with primary emphasis on 1) having a positive impact on the residential course at MIT and potentially residential courses elsewhere, 2) exercising and pushing the boundaries of the online edX infrastructure and platform, and 3) making the content of and experience with the MIT graduate subject more widely accessible worldwide at the level at which it is presented residentially at MIT. To accomplish these goals, the development of 6.341x evolved in four phases on which we elaborate shortly, after first describing the basic organization of the course.

Course organization

Outline

The course 6.341x is an 11-week graduate-level class divided into 18 units. Each unit consists of several topic segments, outlined in Table 6. In a typical week of the spring 2015 MOOC, students were provided with the following:

- *Multiple courseware topics.* Each topic consisted of a combination of brief exercises, text comments, and video segments recorded in a “chalk and talk” format, extracted and edited from lectures of the residential course at MIT, and with slides that were digitally animated specifically for the MOOC. The exercises were interspersed among the video segments specifically to allow students to verify their progress before moving on to the next segment. Consistent with the MIT residential course, lectures contained both mathematically oriented discussions and live demonstrations of signal processing concepts.
- *An overview video from the staff each week.* Videos were recorded in “talking head” style, providing an outline of the week’s topics, in addition to brief, high-level audio signal processing demonstrations illustrating associated concepts.
- *A set of homework problems.* Problems included project-style numerical signal processing problems that students completed using in-browser tools developed specifically for 6.341x.

In addition, three exams were given to evaluate performance. Accompanying the courseware was an online discussion forum on which we comment in more detail below.

Style and format

Elements of the courseware are depicted in Figure 5. Figure 5(b) shows still frames captured from the lecture video segments,

Table 6. MIT 6.341x course outline.

Course Unit	Release Date
Unit 1: Signals and systems in the time and frequency domains	Week 1
Unit 2: Allpass and minimum-phase systems	Week 2
Unit 3: Discrete-time processing of continuous-time signals	Week 2
Unit 4: Sampling rate conversion	Week 3
Unit 5: Quantization and oversampling	Week 3
Unit 6: Signal-flow graph implementations of LCCDEs ^(a)	Week 4
Unit 7: Lattice structures	Week 4
Unit 8: IIR filter design	Week 5
Unit 9: FIR filter design	Week 5
Unit 10: Parametric signal modeling	Week 6
Unit 11: The Levinson recursion	Week 6
Unit 12: Multirate systems and polyphase structures	Week 7
Unit 13: The DFT	Week 8
Unit 14: Computation of the DFT	Week 9
Unit 15: Spectral analysis	Week 10
Unit 16: The TDDTFT ^(b) and modulated filter banks	Week 11
Unit 17: Multirate and critically sampled filter banks	Week 11
Unit E: Enrichment lectures	Weeks 7 and 10

^(a)LCCDEs: Linear constant-coefficient difference equations; ^(b)TDDTFT: time-dependent discrete-time Fourier transform.

featuring chalk and talk-style video clips interspersed with lecture slides animated specifically for 6.341x, as well as in-class signal processing demos. The figure also depicts an example exercise that would fall between video segments, designed for students to verify their understanding. Weekly introductory videos were recorded in a talking head format, shown in Figure 5(a). Figure 5(c) and (d) illustrate interactive elements of two homework problems, in which students were prompted to choose spectral analysis parameters and enter block diagrams.

The basic staff–student interaction model used in the spring 2015 MOOC is depicted in Figure 6. Referring to this figure, the lowest-latency method for staff–student interaction was through the online discussion forum. From the perspective of students, homework problems and exercises provided instant feedback about performance, although aggregate results about student performance were viewable by the staff on a delayed basis. There was also typically a one-week delay between filming and deploying week overview videos, due to the time associated with editing and audio transcription.

Course evolution

The development of 6.341x consisted of four phases that began in the spring of 2013 and continued through the fall of 2015.

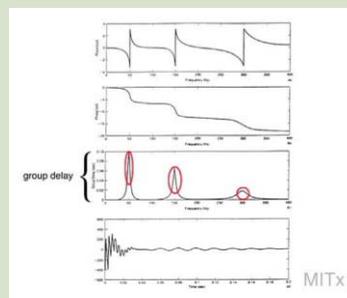
- *Phase I.* Work began in spring 2013 on the translation of significant content of 6.341, as used residentially, into digital form to be used initially as an online augmentation to the residential course. The online platform for digital content was initially Open edX, the open-source platform used residentially by MIT and other schools that mirrors the edX.org infrastructure. It was recognized at the outset that because the residential course covers a graduate-level subject, it would inherently



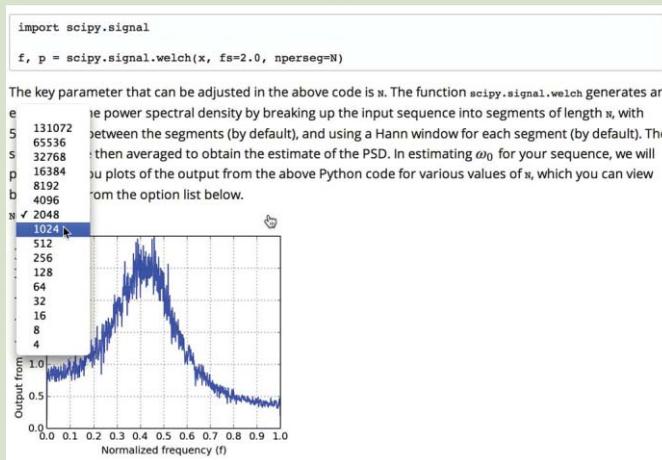
(a)



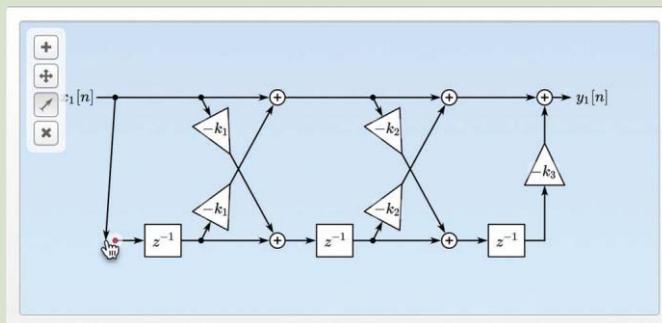
System	Minimum-phase system?
$h[n] = \delta[n - 1]$	<input type="checkbox"/> ?
$y[n] = \frac{1}{2}y[n - 1] + x[n] + 2x[n - 1]$	<input type="checkbox"/> ?
$y[n] = \frac{1}{2}y[n - 1] + x[n] + \frac{1}{2}x[n - 1]$	<input type="checkbox"/> ?
$y(t) = \frac{1}{2} \frac{dy(t)}{dt} + x(t) + \frac{1}{2} \frac{dx(t)}{dt}$	<input type="checkbox"/> ?



(b)



(c)



(d)

FIGURE 5. Screen captures of various elements in the MIT 6.341x courseware. (a) Week overview video with brief audio-based signal processing demonstration. (b) Course topic sequence, composed of in-class lecture videos, interactive problems, and animated slides. (c) The interactive homework problem related to spectral analysis. (d) The interactive portion of homework problem for which students are asked to graphically apply the transposition theorem.

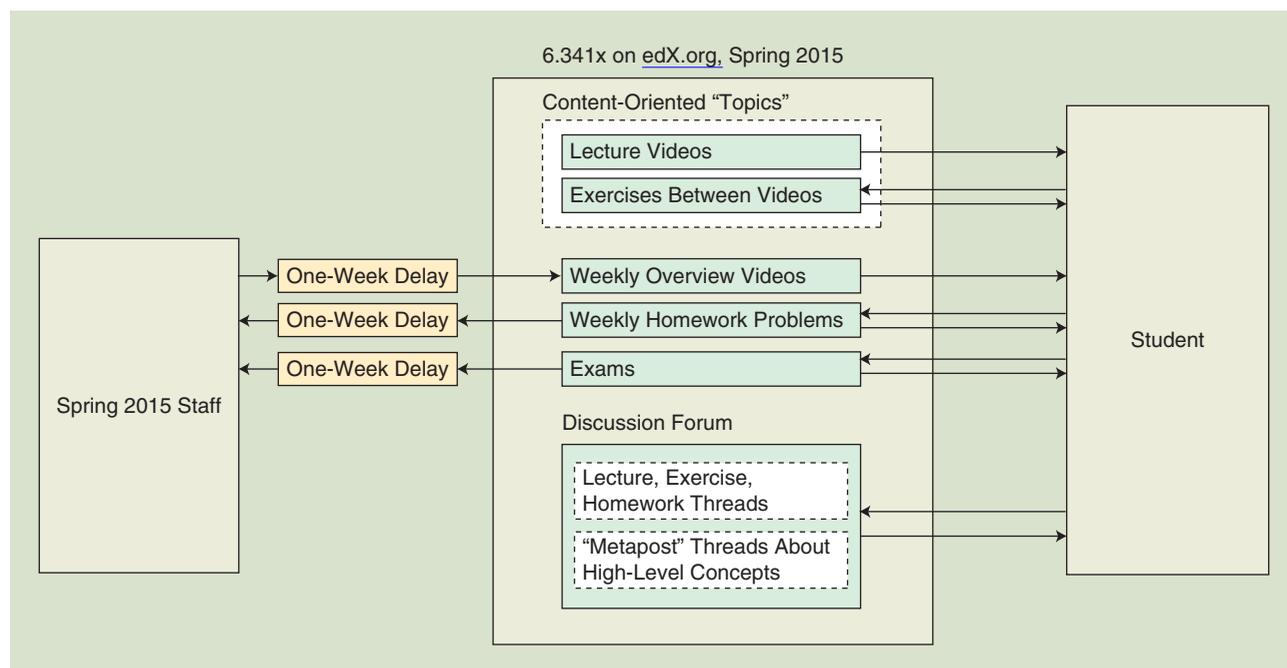


FIGURE 6. A diagram indicating the modes of interaction and flow of information between MIT staff and students, facilitated by the learning platform as it existed in spring 2015.

be more demanding than more introductory online courses. In developing the 6.341x content for this platform, we chose to work “close to the bare metal” of the edX platform to push the boundaries and help to improve the platform.

The process included adapting existing problem sets and a background exam for the online platform, autograding, and creating “finger exercises” to be given to the class and responded to anonymously with fixed-function wireless “clickers” during lecture. A histogram of responses to the exercises, displayed immediately in class, impacted subsequent discussion and pace during the lecture [see Figure 7(b)]. The decision to use fixed-function clickers, as opposed to, e.g., a phone or tablet app, was motivated by the desire to keep the attention of students out of their e-mail and browsers and focused on the lecturer and class discussion. The residential offering of the course in fall 2013 incorporated all of this online content to augment the live lectures. In addition, the live lectures were video recorded.

- **Phase II.** In spring 2014, the video recordings of fall 2013 lectures were heavily edited for crisper pacing, generating video segments of appropriate content and length, and sequencing with the finger exercises. In the residential offering, fall 2013 students were given a background exam on Open edX and completed autograded problem sets on the platform. The problem sets also used an in-browser “explanation box” system that was developed for the residential course, in which students could enter symbolic equations, proofs, and reasoning about their answers [see Figure 7(a)]. These were then electronically distributed to the course staff for manual grading. Given the emphasis on a deep conceptual understanding that has traditionally been a key part of the

residential course, the comment box system provided a way to assess this understanding without diluting the problems to fit within the constraints of an autograding system.

The success of the online experience in the residential course was a key motivation for carrying the material forward to a MOOC. Toward this end, the digital content was continually expanded and refined until it was ready to be run as a private beta for a limited number of participants from industry in fall 2014, and then publicly deployed on edX.org in spring 2015. This transition was a significant effort, funded largely by the MIT electrical engineering and computer science (EECS) department and the MIT Office of Digital Learning (ODL), together with edX. A significant effort was required to edit the in-class video recordings for the MOOC environment, which we found works best with short, ten-to 20-minute, well-paced segments. The in-class exercises were also augmented and in some cases modified for a better match to the edX environment.

- **Phase III.** This phase consisted of first offering 6.341x as a MOOC in fall 2014 in a beta version limited to 200 participants from industry, and then in spring 2015 as a fully open online course. For all registered students, the autograded online background exam was made available before the start of the MOOC to allow participants to assess their background relative to the course content.
- **Phase IV.** In the fall semester of 2015, residential 6.341 was offered at MIT with the usual structure of three hours of live class time per week, one hour of live recitation discussion with the teaching assistants (TAs), and a handwritten midterm and final exam graded by the staff. The course made full use of all of the digital online 6.341x content

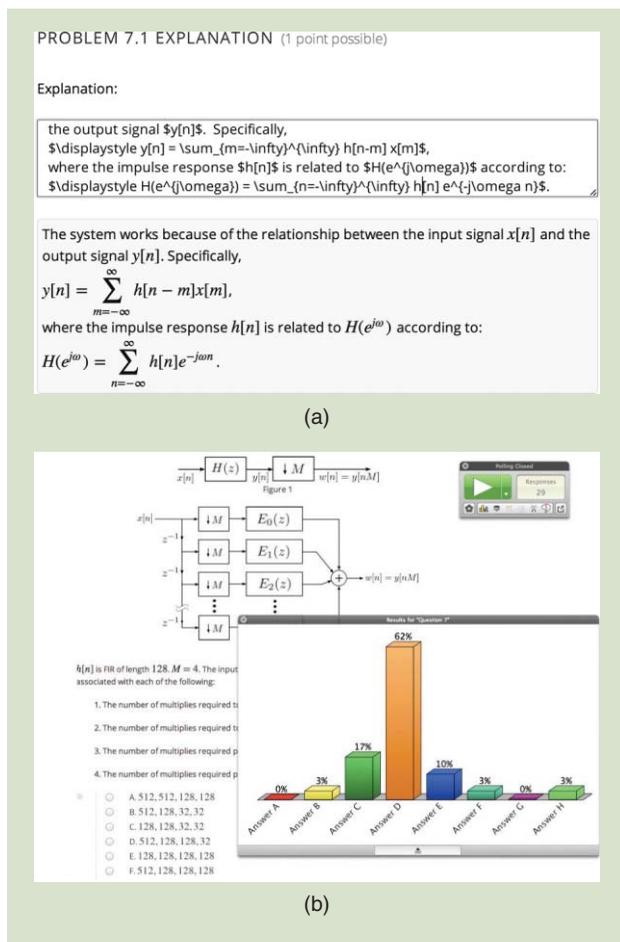


FIGURE 7. Additional online tools used specifically for residential deployment of MIT 6.341x. (a) The MathJax-enabled student “explanation box,” in which residential students provided staff with reasoning and derivations behind their autograded answers. (b) A polling system used with in-class “finger exercises,” designed to provide self-assessment of student understanding during lectures.

running on the MITx platform. Participants were asked and strongly encouraged to watch the videos and work the finger exercises before coming to the live class sessions. Class sessions were structured to give an overview and incorporate many subtleties about the topic under the assumption that students had previewed the content online. During the class sessions, a number of finger exercises on the topic were presented as clicker questions with instant feedback to the staff and class through display of the histogram of responses. The syllabus schedule was also structured under the assumption that the students would preview the topics online before class sessions. Consequently, online content for week N was made available online to the class late Thursday evening in week $N - 1$. All homework on the topics in week N was due at the end of that week and was autograded online, but each problem also included an explanation box in which the student could elaborate on the solution. The staff reviewed these explanation boxes and student feedback was given quickly. Staff solutions were posted online immediately after the due date.

The overall experience for both students and staff in incorporating all of the digital online content and emphasizing the importance of previewing in both the structure of the schedule and in the level and pacing of the course was extremely positive.

Numerical examples and tools

In the residential course, multiple numerically focused class projects are traditionally assigned to students to provide practice with the use of IIR and finite impulse response (FIR) filter design and order estimation tools and with the use of various methods for spectral estimation. These projects have been very well received by students taking the residential courses at MIT. For the MOOC, we attempted to replicate or at least approximate the experience by creating a set of in-browser tools that learners were able to use to perform many of the numerical tasks associated with these projects.

As one example, we created a set of in-browser tools, discussed in greater detail in “Numerical Tools for Filter Design (MIT),” for performing the numerical tasks commonly associated with FIR and IIR filter design. The tools allowed learners to estimate design parameters, compute filter coefficients, and compare the performance of the resulting filters against stated specifications. The numerical output from the tools was then evaluated in the context of various assigned problems, which provided textual as well as graphical feedback about the submitted filter designs. For example, any regions of the magnitude response violating the stated constraints would automatically be highlighted, indicating to the learner where specifications had been violated. Automated textual messages about the submitted designs were also programmed to respond to common pitfalls that we identified as having commonly occurred in class projects during past semesters of the residential course.

Personnel

Through its four phases, the development of 6.341x required considerable resources and support from the MIT EECS department and ODL and the edX team. The responsibility for content development and for incorporation of the content into the online platform was ours along with Tarek Lahlou, an EECS graduate student, who also was an instructor for the industrial beta version of the MOOC. In developing interactive content and incorporating it into the online platform, we collaborated closely with the edX and ODL teams and, in particular, with TC Haldi, Tsinu Hermano, Joe Martis, and Peter Pinch.

A major effort in developing content was required for editing and reformatting the in-class live video recordings into segments with good pacing and length. This editing was the responsibility of Alan Oppenheim, Tom Baran, and Isaac Chuang, together with video editors Jim Ohm and Edwin Cabrera. The video segments were also further reviewed for accuracy and appropriate highlighting by Tarek Lahlou, as well as EECS graduate students Anuran Makur and Lucas

Numerical Tools for Filter Design (MIT)

In 6.341 at the Massachusetts Institute of Technology (MIT), a numerically focused class project on IIR and FIR filter design was traditionally assigned as a key component of the course. In developing 6.341x, one of our goals was to provide a project in this theme for online learners. With this in mind, we wanted to deliver a numerical project experience that was contained entirely in the browser, provided graphical and textual feedback to the learner about how and where their numerical input might be incorrect, and did not require the learner to have extensive knowledge of a particular numerical signal processing language, yet still provided practice dealing with many of the numerical issues associated with using such packages.

With these goals in mind, our approach in 6.341x was to write a series of server-side Python libraries that allowed learners to perform order estimation and filter design by submitting design parameters to the edX server, which would then return the numerical output that was computed from the parameters. Learners were also provided with associated Python/SciPy code to reproduce these results on their own machine if they wished, although very little knowledge of Python was required to use the online tools.

Using the numerical designs obtained from the in-browser tools, learners were able to complete various assessment problems, and basic code used in performing the assessment was provided to learners to use as a reference if they wished. Graphical feedback based on learner input was generated dynamically on the server and passed to the browser, e.g., highlighting regions of the magnitude response where a numerical design might not have met the stated specifications.

From the perspective of the student, the process of designing a particular filter typically involved the following sequence. First, an order estimation tool would be used, as

depicted in Figure S3. The returned values could then be used to select parameters in a corresponding filter design tool, shown in Figure S4. By transferring the resulting numerical design to the appropriate assessment problem, the learner was provided with automated, graphical feedback about his or her design, as is depicted in Figure S5.

The overall reaction to the exercises using these tools was positive. Learners indicated, in particular, that the freedom to explore various design methods and parameter choices was a key part of the learning experience, highlighting to us the value of providing access to numerical tools and problems in an online signal processing course.

You can use this tool to design a Parks-McClellan lowpass filter using pre-specified design parameters. These parameters can be selected any number of ways, including using the order estimation tool above.

Using this tool is equivalent to executing the following in Python, assuming that `scipy.signal` is installed. Similar syntax can also be used in Matlab and GNU Octave.

```
import scipy.signal
b = scipy.signal.remez(numtaps, bands, amps, weights, Hz=2.0)
```

The above code designs a Parks-McClellan filter and returns the design as a vector of polynomial coefficients `b`, i.e. as a list containing the impulse response of the filter.

Enter your input parameters below, and select "Check" to view the computed output.

numtaps =

bands =

amps =

weights =

Output parameters:

```
b = [-2.70574817e-04, 2.12148120e-03, 3.42001684e-03, 5.53334447e-03, 7.91225078e-03, 1.01323744e-02, 1.16873955e-02, 1.20422095e-02, 1.07332530e-02, 7.49672913e-03, 2.37311229e-03, -4.21249419e-03, -1.14380070e-02, -1.81351657e-02, -2.29212983e-02, -2.43867033e-02, -2.13298901e-02, -1.29798901e-02, 8.14028458e-04, 1.94792118e-02, 4.17101386e-02, 6.55886051e-02, 8.87933807e-02, 1.08906364e-01, 1.23728885e-01, 1.31587168e-01, 1.31587168e-01, 1.23728885e-01, 1.08906364e-01, 8.87933807e-02, 6.55886051e-02, 4.17101386e-02, 1.94792118e-02, 8.14028458e-04, -1.29798901e-02, -2.13298901e-02, -2.43867033e-02, -2.29212983e-02, -1.81351657e-02, -1.14380070e-02, -4.21249419e-03, 2.37311229e-03, 7.49672913e-03, 1.07332530e-02, 1.20422095e-02, 1.16873955e-02, 1.01323744e-02, 7.91225078e-03, 5.53334447e-03, 3.42001684e-03, 2.12148120e-03, -2.70574817e-04]
```

FIGURE S4. An in-browser tool for performing minimax-optimal FIR filter design.

You can use this tool to estimate the parameters required to design a Parks-McClellan lowpass filter. It is equivalent to executing the following in Python, assuming that `dtsp.py` is installed. Similar syntax can also be used in Matlab and GNU Octave.

```
import dtsp
numtaps, bands, amps, weights = dtsp.remezord(wp/2.0, ws/2.0, [1, 0], [dpass, dstop], Hz=1.0)
bands *= 2.0 # above function outputs frequencies normalized from 0.0 to 0.5
```

Enter your input parameters below, and select "Check" to view the computed output.

Passband edge frequency (normalized from 0 to 1): $w_p =$

Stopband edge frequency (normalized from 0 to 1): $w_s =$

Passband ripple about unity (linear): $d_{pass} =$

Stopband ripple about zero (linear): $d_{stop} =$

Output parameters:

```
numtaps = 52
bands = [0.00000000e+00, 1.07140000e-01, 1.78570000e-01, 1.00000000e+00]
amps = [1.00000000e+00, 0.00000000e+00]
weights = [1.00000000e+00, 3.33333333e+01]
```

FIGURE S3. An in-browser tool for performing FIR filter order estimation.

Enter coefficients for a Parks-McClellan filter that meets the above stated specifications (1-7). Specify the impulse response $h[n]$ of your filter using the parameters in Eq. 6.1.

$k =$

$[b_0, \dots, b_N] =$

From the number of elements in $[b_0, \dots, b_N]$, your value of N was inferred to be $N = 51$.

Issue: The staff believes that a larger value of N is required.

Issue: Your filter did not meet the magnitude response specifications:

- The maximum passband gain was exceeded.

The magnitude response of your filter, obtained numerically from your input, is depicted below in logarithmic and linear units, respectively emphasizing the stopband and passband performance.

FIGURE S5. An exercise assessing a particular filter design entered by a learner, providing automated feedback about where the design exceeded specifications stated previously in the problem.

The Editor: A Browser-Based Tool for Manipulating Signal-Flow Diagrams (MIT)

In the residential course at MIT, on which 6.341x was based, a key focus was traditionally placed on using both symbolic and numerical exercises in assessing understanding and many times, in particular, using signal-flow and block diagrams in doing so. With these goals in mind, we wrote a variety of assessment exercises around what we referred to as “The Editor”: an in-browser graphical tool that we built, designed specifically to give learners the ability to create and manipulate numerical and symbolic signal-flow diagrams directly inside the courseware.

A causal, LTI system is depicted in the following block diagram.

Modify the multiplier coefficients and feed-forward delay in the block diagram so that the system has the following impulse response from $x[n]$ to $y[n]$:

$$h[n] = \delta[n] + 2\delta[n - 1] + 4\delta[n - 2] + 8\delta[n - 3].$$

The system in the block diagram is assumed to be causal and LTI for any set of parameters that you may choose.

FIGURE S6. A problem for which learners specify block diagram parameters using The Editor.

The Editor is a JavaScript library that couples a declarative representation of a signal-flow structure with a graphical interface in which the representation can be manipulated. The signal processing representation (SPR) is based on extensible markup language (XML; referred to as SPRXML) and encodes the topology and parameters associated with a particular signal-flow block diagram. Using associated server-side libraries also written for the course, the edX server can dynamically generate an SPRXML system, pass the system to the browser where it is displayed

The following system contains an LTI subsystem whose z-transform is written as $H(z^L)$, with L being a positive integer.

Apply the noble identity to this system by performing manipulations in the Editor window below, i.e. so that $y_1[n] = y_2[n]$ when $x_1[n] = x_2[n]$.

FIGURE S7. A problem assessing the application of the noble identity by performing manipulations involving symbolic expressions using The Editor.

Nissenbaum. Further behind the scenes are many others at edX and ODL, without whose active involvement and support the interactive numerical content would not have been able to run on the platform.

Unique elements

Discussion forum

A key component of 6.341x was the edX discussion forum that was very actively monitored by the course instructors (Tom Baran and Alan Oppenheim for the public MOOC, and Tarek Lahlou for the private industry beta), as well as by several community TAs (CTAs) from industry, who had participated in the limited beta run in fall 2014. On the forum, students engaged with one another and the staff, discussing course content and how it might be applied to their own engineering problems. By the end of the spring 2015 MOOC, a lively community of engineering professionals, students, independent learners, and educators had emerged on the 6.341x discussion forum.

As indicated in Figure 6, the online discussion forum was the lowest-latency mode of student–staff interaction

and, as such, was very actively used throughout the duration of the course. There were typically two types of content-oriented posts on the forum: specific questions about homework problems (typically generated by students) and regular more-elaborate posts written by the staff, designed as a springboard for higher-level discussion about various signal processing concepts. For the spring 2015 MOOC, a total of six CTAs were also available on the forum, selected from those students who performed well previously in the industry beta version.

Based on course feedback, students generally felt 6.341x staff to be very accessible via the online forum. The staff regularly monitored the progress of threads on the forum and encouraged discussion among the students, e.g., by posting comments and follow-up questions. Staff responses to questions about homework problems were intentionally delayed somewhat unless an error had been identified, giving students the opportunity to respond first and further encouraging students to view the forum as a collaborative meeting place among a community of learners, as opposed to as a resource for homework help from the staff.

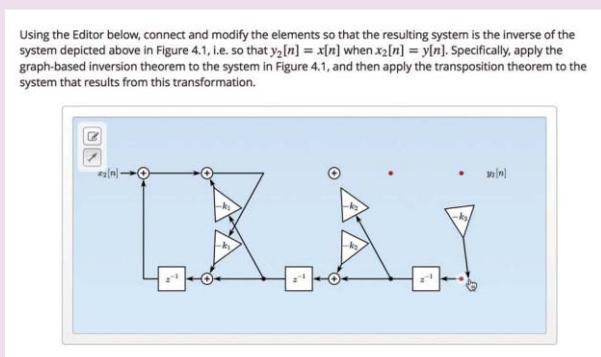


FIGURE S8. An Editor-based problem assessing the detailed application of graph theorems.

by The Editor, and after the learner manipulates the system graphically, the server grades the modified SPRXML.

In 6.341x, a variety of problem types were written around The Editor. For example, The Editor was used in several problems to display a dynamically generated block diagram having parameters that were updated on each attempt. This was used by learners who desired repeated practice in computing transfer functions. Learners could also use The Editor to modify parameters in a block diagram having a fixed topology, as shown in Figure S6. By using the LaTeX-like equation syntax supported by The Editor, in conjunction with a symbolic grader, the course was also able to assess the ability of a learner to apply key signal processing identities, such as the noble identity as depicted in Figure S7. Detailed assessments of the use of identities were also possible using

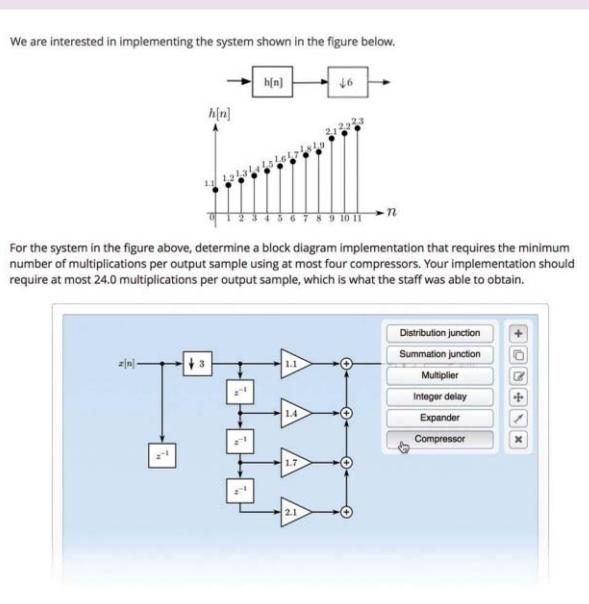


FIGURE S9. A system implementation exercise in which learners use The Editor to specify an efficient multirate system.

The Editor, such as in the problem shown in Figure S8, which assesses the application of the transposition and graph-based inversion theorems. Using The Editor, block diagrams could also be created from scratch. The problem depicted in Figure S9 uses The Editor in assessing the ability of a learner to create an efficient implementation of a multirate system.

Platform augmentation and development

A key goal in developing 6.341x was to design content that pushed the boundaries of and extended the edX platform. This goal was made in collaboration with the MIT ODL and edX and was done for several reasons. For example, we considered it important to provide online access to numerical tools for students who might not have had specific signal processing packages available on their machine (e.g., MATLAB, LabView, etc.). Another key reason was that material in 6.341x is intended to teach concepts and not specific languages. This has historically been true, even with those problems and projects in the residential course that are numerically focused. With this in mind, on the 6.341x site, various numerical tools were provided. Parameters could be entered into the browser, the input would be evaluated on the server, and in addition to providing the numerical result, Python code would be provided so that students could deploy the result on their own system, if they wished. An example of the workflow associated with using these tools in the context of FIR and IIR filter design is depicted in Figures S7–S9 and discussed in “Numerical Tools for Filter Design (MIT).”

Another tool developed for 6.341x is what became known as The Editor, a JavaScript-based interface for graphically entering symbolic block diagrams that was used in a variety of problems and exercises. The use of The Editor is depicted in Figure 5(d) and Figures S3–S6, and is discussed in greater detail in “The Editor: A Browser-Based Tool for Manipulating Signal-Flow Diagrams (MIT).” Using The Editor, a student could enter systems composed of standard signal processing blocks such as summation nodes, coefficient multipliers, expanders, and decimators. Unlike various traditional signal-flow entry tools, The Editor supported the entry of blocks having symbolic parameters. A learner could be asked, for example, to symbolically apply a noble identity or transposition theorem to a preloaded signal-flow system, and the auto-grader would symbolically evaluate whether the properties had been correctly applied.

There were also problems in which learners used The Editor to input block diagrams from scratch, implementing, for example, a multirate system having a desired response. In this case, the entered block diagram would be scheduled and implemented on the edX server, with the

resulting automatically generated implementation used as a basis for evaluation. Motivated by the goal of providing detailed feedback to students, the scheduling algorithm was also able to automatically analyze and reduce algebraic loops, forming an online implementation of the algorithm as shown in [16].

Class by the numbers

Retention and engagement

In the spring 2015 run of 6.341x, the course began with approximately 10,500 learners registered and closed with approximately 9,500 learners registered and 110 receiving certificates. Early in the course, about 2,500 registrants were clearly active, with approximately 500 active at the close of the course. Perhaps not surprisingly, the largest spikes and subsequent drop offs in the number of registrants that were active coincided with the three exams.

Demographics

The residential course 6.341 is taught once each academic year at MIT, with typical end-of-semester enrollment of 30–40 (mostly) graduate students. With the acceptance rate of 2–3% in the MIT EECS graduate program, it is reasonable to assume that the students in the residential course have strong backgrounds and are well qualified. Attrition rates in 6.341 are typically on the order of 30% between the first and last weeks of the semester.

In the industry beta version of 6.341x, there were a total of approximately 170 registered students, with 36 completing the course with a passing grade (i.e., performance at the level of an A or B grade). In the spring 2015 open run of the MOOC, the initial registrants represented a total of 136 countries. Among those who successfully completed the spring 2015 MOOC, the overwhelming majority were those who joined from industry. Keeping in mind our goal of reaching a broader number of individuals than we were able to reach residentially at MIT, we were delighted that while maintaining the same level of content depth, difficulty, and sophistication, we were able in a single run of the course to impact the same number of students as would be impacted residentially during approximately six–seven years of teaching at MIT. In addition, as further evidenced by the participant quotes and the modest percentage of participants able to successfully complete the course, it was difficult and demanding in terms of time and background, as would be expected with a course at the graduate level. It was also especially encouraging that a significant percentage of those impacted by the MOOC were individuals who do not traditionally constitute a major component of the MIT EECS graduate-level student demographic; specifically, those completing 6.341x were primarily university

faculty and senior-level engineers working on projects at well-known high-tech firms.

Impact on the MIT residential course

We have a very clear and strong sense of how MOOC content has enhanced the residential experience and its impact for the future. Although MOOC content by its nature is designed to stand alone, in the residential context it becomes a very strong multiplier on the in-class time that students have with the staff. In the residential course, students were strongly encouraged to preview the online content before coming to class, and the assumption during class time was that students had previewed the videos and worked at least some of the online exercises. In the context of a graduate-level course that carries an assumption of a high level of maturity and commitment on the part of students, we chose not to make previewing explicitly mandatory, nor to have a graded mini-quiz at the beginning of each in-class session. Furthermore, class sessions were not simply discussions or Q&A sessions. However, the level and pace of the presentation was predicated on the assumption of previewing. The approximate experience was that about one-third of the class regularly previewed the topics, one-third sporadically or superficially did so, and about one-third almost never did at all. Those who did uniformly (anecdotally) felt that it made a significant difference. And whether or not a student regularly previewed the content, virtually all actively used the online content after in-class sessions.

Our conclusion is that the residential course benefited enormously from the availability of the rich online content and that students actively used it both before and after the in-class interaction. The MOOC by itself is no substitute for a well-taught residential course. However, it can be a significant

enhancement to any residential course, and in our view it is a strong substitute for any poorly taught residential course.

enhancement to any residential course, and in our view it is a strong substitute for any poorly taught residential course.

Feedback

Overall feedback about 6.341x was strongly positive, and perhaps the most common negative feedback was that the level of sophistication and time commitment required to take 6.341x was higher than expected. However, for those who met the background prerequisites, this was viewed as an asset, commenting that 6.341x stood apart from other MOOCs in this regard. Those students whose background was slightly weak but who actively engaged with the forum generally found that the availability and encouragement of course staff and CTAs allowed them to brush up on their weak spots and stay engaged with the course.

As expressed by one of the students, “Right before the course started, I thought I was well prepared for this course. After all, I have a strong background in [signals and systems], and I am very familiar with digital signal processing. But after the first few weeks...I found the exercises and problem sets challenging... I almost gave up [were it not for] the helpful feedback from

Much in the same way, the openness of online courses often causes prospective learners to underestimate the importance of prerequisites, which leads to the low yield rates that most other instructors experienced.

instructors, TAs, and other kind peers.” Regarding the forum, a student commented, “The instructors did a fantastic job in interacting with the students. I cannot recall one single question that was not properly addressed. They are very kind and responsive.” To us, this indicated that better learning outcomes in 6.341x were strongly facilitated by a high level of staff involvement on the forum.

In terms of the organization of the courseware, students reacted especially positively to the organization of the course topics, verifying the staff’s intuition that weaving video segments with short exercises would provide a natural mechanism for students to check their understanding before moving too far along. The interactive numerical content in the courseware was well received and generally viewed as a unique feature of 6.341x. Some students had been hoping for an opportunity to write and test their own signal processing code, which fell outside of the scope of the interactive problems. Many of those students desiring additional code practice were typically sufficiently motivated to do so on their own and post the results of their efforts on the discussion forum, significantly enriching the discussion among the community of learners.

A number of individuals also commented that, overall, 6.341x had a significant impact on them both personally and professionally. In particular, students indicated the immediate applicability of 6.341x to their professional work in fields ranging from software-defined radio to the design of particle accelerators. A recurrent theme in student feedback was also that 6.341x helped them to decide to change their field to signal processing, which the staff was delighted to hear. Overall, these comments indicate to us that 6.341x is a challenging course, but for those who complete it, it is also a very rewarding experience.

Next steps

Going forward, there are several potential modes and roles for 6.341x content. In the context of its incorporation into the MIT residential course, there is no question that it has contributed significantly to the educational experience, and it will continue to be incorporated and developed further in the context of the MIT residential graduate course. We welcome the opportunity for it to find a similar role in many residential courses at this level elsewhere.

In the form that 6.341x ran as a MOOC in spring 2015, the content was released on a fixed schedule, and learners were expected to commit to that schedule. Although that mode of delivering difficult course content of this depth and sophistication is typical in a university environment, rigid pacing is not necessarily well matched to participants outside of the university environment or to those with other significant time constraints and deadlines. In future deployments of 6.341x, we envision a self-paced mode on a more local platform so that pacing of the content can accommodate the needs of particular groups of participants. As one model for use in an industry environment, a company could perhaps subscribe to the content and platform for use internally. The oversight, pacing, and staffing of the course could then be managed internally to match the needs and schedule constraints of participants.

Our current plan in the near term is to release the total 6.341x course content in a form that is freely available to

learners on the edX platform, for the purpose of self-study. In this mode, autograding the exercises and problem sets will be activated, but no discussion forum or support staff will be available. Specifically, it will be accessible in a manner similar to that of course content on MIT OCW, although with a more interactive component. When it becomes available, the material will be accessible at <http://www.rle.mit.edu/dspg/6.341x>.

Conclusions

The experience of the three DSP courses presented in this article clearly suggests that online platforms and content offer rich opportunities for teaching signal processing. How to best affect this is not yet clear, and “best practices” can be very dependent on the demographics of the learners and the objectives and personal style and preferences of the course developers and instructors in adapting residential course content to an online environment.

Commonalities

Perhaps the most important commonality across the three courses is the focus on solid theoretical foundations. In that and a number of other respects, they have a different purpose and target audience than many other online courses, for which content is primarily oriented toward a high-level overview of a topic area. This difference is clearly a key factor in the drop-off level of active involvement from registration (i.e., many registrants are motivated mainly by curiosity) to course completion. This is inevitable for online courses that attempt, to the extent possible, to provide participants with the same depth and sophistication as a residential course. It is also important to recognize that a MOOC is no substitute for a well-taught residential course that incorporates significant interactive face time with a knowledgeable and motivated staff.

Because the three courses are based on residential classes at different levels, the backgrounds and expectations of the participants somewhat differ. However, in a broad sense, a serious background in signals and systems at some level was common, and quite often, a more advanced background including industrial project experience with a partial motivation to refresh that background was helpful.

Differences

Two of the courses were offered on edX and one on Coursera. The differences between platforms are certainly many but not profound enough to significantly affect the way the material was structured and presented. The three courses were, in fact, more distinct in their handling of numerical exercises and examples. EPFL ultimately gravitated toward Python (via IPython Notebooks), Rice experimented with a tight integration between MATLAB and edX, and MIT developed specific extensions to the edX platform to provide in-browser numerical exercises independent of any specific package or programming language.

Each of these approaches has potential advantages and drawbacks. MATLAB offers perhaps the most complete signal processing sandbox and a very user-friendly learning curve, but its scripting language does not please those students with a more rigorous background in computer science.

Furthermore, it is a commercial solution, and although free alternatives do exist, they are not as complete and robust. As an alternative, Python is becoming increasingly popular in scientific programming, but the language itself is still embroiled in a difficult version transition. IPython Notebooks are a very versatile didactic tool, but they do not scale well to large projects and do not offer easy version control. Adhoc code and browser extensions are clearly the most attractive approach with respect to integration with lectures and from the point of view of user experience; but, inevitably, they impose a very high development and maintenance cost on teaching staff.

Lessons learned from our MOOC experiences

- A very clear and strong lesson we all learned is that the resources needed in terms of effort, financing, and platform backup to successfully develop and run an online course with serious depth and content are enormous. Video segments need to be short and crisp and, even if extracted from in-class video recordings, major editing is essential. Exercises, problems, and projects all need to be carefully designed and restructured, even when based on residential course content.
- The backgrounds of MOOC participants are typically very diverse. For our courses, they included educators, experienced engineers, high school and college students, and retirees. Making clear to potential participants the assumed background required and providing a preliminary background exam for their calibration before registering would seem essential.
- In stark contrast to a residential setting using MOOC and other online resources, in a MOOC setting, it is extremely difficult to exercise and test deep understanding of concepts.
- The opportunity for students to receive immediate feedback as they work through exercises, problems, and projects is a key defining feature of teaching when using online resources, whether residentially or as MOOCs.
- With the large number of participants in a MOOC and the analytics that the platforms can capture and provide, there are unprecedented amounts of data on what works and what can be improved; these data are in the form of direct feedback from the students on the forums, indirect observation of self-regulating conversations among students, performance on the various elements, and information on the use of the videos. MOOC platforms in use today log every interaction between the learner and the interface: timestamps, number of views, fragmentation of video consumption, and access to previous material are just a few of the variables to which we now have access. For now, these data remain largely untapped. Clearly, the insight from such a vast data set would benefit not only online teaching but residential courses as well.
- The emergence of volunteer CTAs is one of the things that makes MOOCs truly different from usual residential courses. It illustrates that educational communities have the potential for emergent behavior, where students mentor and tutor each other with little interaction from an instructor. Moving

forward, it is important to find ways to incentivize and support this very positive and useful behavior.

The role of certification

It seems clear anecdotally that a high percentage of MOOC participants value some form of certification of their successful completion of the course. What we find to be less clear is the inherent value of a statement of accomplishment, how it might fit into a student's curriculum, and how a professional can leverage its value in the workplace. Our institutions have been very prudent in the wording of certificates and have made sure to prevent any association between MOOCs and the actual on-campus curriculum. Clearly, any other course of action would be difficult in the absence of a reliable method to assess the identity (and the proficiency) of online students. If it is already hard to manage the test administration process on campus, the difficulties online are close to insurmountable: Multiple identities are easy to forge, cheating is easy, and unless exams are constantly rewritten, solutions from previous editions of the class are just a click away. In this sense, the "massive" and "open" characteristics of MOOCs are also their liability as far as proper certification is concerned. Much in the same way, the openness of online courses often causes prospective learners to underestimate the importance of prerequisites, which leads to the low yield rates that we (and most other instructors) experienced. Once again, it will be difficult to arrive at a compromise for which the original spirit of the MOOC "revolution" coexists with a preselection process.

The big picture, with an eye on the future

We started by evoking "the year of the MOOC" and conclude with some reflections on the future of MOOCs, given our collective experience so far. Perhaps the first observation is that, as is widely recognized in the community, a clear business model for making MOOCs financially viable and sustainable has yet to emerge. Online courses require a huge amount of work to design, realize, and sustain, which adds up to a significant financial investment on the part of the sponsoring institutions. Yet completion rates are so low that any residential class with similar drop-off rates would be unsustainable. Potential solutions lie in directions such as the evolution toward specialization classes with fewer and prescreened participants and/or the targeted professional market. Such directions, of course, are no longer "massively open" but continue to take advantage of the enormous benefits of the online environment. Perhaps the harsher realization is that a deep understanding of a topic is built from a solid foundational background and then serious and hard work to advance that background into a deeper and richer understanding. Absorbing difficult content is, well, difficult, no matter which delivery channel is used to reach the students.

These difficulties should not overshadow the enormous potential of MOOCs and the content that they contain. They are typically based on packaging extremely high-quality material from on-campus classes into an attractive format accessible by anyone from anywhere. There is a strong similarity to the process of evolving course notes, often hastily typed and poorly

photocopied, into high-quality textbooks that then impact a much wider audience. And, of course, the significant effort that goes into polishing course notes into widely available textbooks, often underestimated, has a major impact on the residential teaching of that content, both locally and more broadly. Use of online MOOC content to enhance residential teaching of the material appears to have enormous potential and, as with well-written textbooks, can provide enormous leverage to dedicated residential courses worldwide. However, there is often the misconception that incorporation of MOOC content and other online content will lead to cost savings and reduced required effort. In our experience, this is not the case if teaching quality remains important. Incorporation of this content into residential courses has the potential for enormous leveraging and enhancement to materially increase the quality of the education. These are indeed exciting times for education!

Authors

Thomas A. Baran (tbaran@alum.mit.edu) received his M.S. and Ph.D. degrees from the Massachusetts Institute of Technology (MIT). He has been a visiting lecturer at MIT and edX (MITx) fellow at MIT and is a cofounder and chief executive officer of Lumii, Inc. His research interests include the intersection of signal processing with optimization theory as well as light-field signal processing and audio signal processing. He has received awards from the IEEE and MIT in recognition of research, teaching, and mentorship.

Richard G. Baraniuk (richb@ece.rice.edu) is the Cameron Professor of Electrical and Computer Engineering at Rice University in Houston, Texas. He received the IEEE Signal Processing Society Best Paper Award, Best Column Award, Education Award, and Technical Achievement Award. For his education projects, including Connexions and OpenStax, he received the National Outstanding Teaching Award from Eta Kappa Nu, the Tech Museum Laureate Award, the Harvard Law School Internet Pioneer Award, the World Technology Education Award, and the IEEE James H. Mulligan, Jr. Medal for Education. His research interests include signal processing, machine learning, and open education. He is a Fellow of the IEEE and the American Association for the Advancement of Science.

Alan V. Oppenheim (avo@mit.edu) is Ford Professor of Engineering in the Electrical Engineering and Computer Science Department at the Massachusetts Institute of Technology (MIT). He is a coauthor of a number of widely used textbooks on signal processing. He received an honorary doctorate from Tel Aviv University and was a Guggenheim fellow and a Sackler fellow. He received multiple IEEE awards, IEEE Signal Processing Society awards, and MIT awards in recognition of outstanding research, teaching, and mentoring. He is a member of the National Academy of Engineering. He is a Life Fellow of the IEEE.

Paolo Prandoni (paolo.prandoni@epfl.ch) is a lecturer at the École Polytechnique Fédérale de Lausanne, Switzerland. He coauthored the textbook *Signal Processing for Communication* and designed the first digital signal processing massive open online course for Coursera in 2012. He is a

cofounder and chief science author of Quividi, a leading video analytics company for out-of-home audience measurement.

Martin Vetterli (martin.vetterli@epfl.ch) received his Dipl.-Ing degree from the Swiss Federal Institute of Technology, Zürich, in 1981; his M.S. degree from Stanford University, California, in 1982; and his Ph.D. degree from the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, in 1986. He held faculty positions at Columbia University, the University of California, Berkeley, and EPFL, where he is a professor. Since 2013, he has been the president of the Swiss National Science Foundation. His research interests include electrical engineering, computer sciences, and applied mathematics. He is the coauthor of three textbooks and numerous papers and patents and received several awards for his research. He is a fellow of the Association for Computing Machinery and a foreign member of the National Academy of Engineering. He is a Fellow of the IEEE.

References

- [1] T. Friedman, Revolution Hits the Universities, *The New York Times*, Jan. 26, 2013.
- [2] A. V. Oppenheim, "Videotape lecture series and study guide on digital signal processing," MIT Center for Advanced Engineering Studies, Sept. 1975. [Online]. Available: <http://ocw.mit.edu/resources/res-6-008-digital-signal-processing-spring-2011/>
- [3] A. V. Oppenheim. (1987, Apr.). Videotape lecture series and study guide on signals and systems: An introduction to analog and digital signal processing. *MIT Center for Advanced Engineering Studies*. [Online]. Available: <http://ocw.mit.edu/resources/res-6-007-signals-and-systems-spring-2011/>
- [4] P. Prandoni and M. Vetterli, *Signal Processing for Communications*, 2008. [Online]. Available: <http://www.sp4comm.org>
- [5] [Online]. Available: <https://www.ctan.org/pkg/dspricks>
- [6] H. J. Trussell and D. Baron, "Creating analytic online homework for digital signal processing," *IEEE Signal Processing Mag.*, vol. 32, no. 5, pp. 112–118, Sept. 2015.
- [7] [Online]. Available: <http://ipython.org/notebook.html>
- [8] [Online]. Available: <https://github.com/LCAV/SignalsOfTheDay>.
- [9] M. Martinez-Camara et al., "The Fukushima Inverse problem," in *Proc. 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, British Columbia, Canada, pp. 4330–4334.
- [10] I. Dokmani, "Listening to distances and hearing shapes: Inverse problems in room acoustics and beyond," Ph.D. dissertation, EPFL, Lausanne, Switzerland, 2015.
- [11] [Online]. Available: https://en.wikipedia.org/wiki/Johann_Heinrich_Pestalozzi
- [12] [Online]. Available <http://www.katyjordan.com/MOOCproject.html>
- [13] [Online]. Available: <http://openstax.org>
- [14] R. G. Baraniuk. (2016). ELEC301.2x: Discrete-time signals and systems, part 2: Frequency domain. Rice Univ. [Online]. Available: <https://courses.edx.org/courses/RiceX/ELEC301.2x/2015Q3/info>
- [15] A. V. Oppenheim and T. A. Baran. (2015). 6.341x Discrete-time signal processing. edX. [Online]. Available: <https://www.edx.org/course/mitx/mitx-6-341x-discrete-time-signal-4396>
- [16] T. A. Baran and T. A. Lahlou, "Implementation of interconnective systems," in *Proc. 2015 IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, South Brisbane, Australia, pp. 1101–1105.
- [17] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 2010.
- [18] A. V. Oppenheim and G. Verghese, *Signals, Systems and Inference*. Englewood Cliffs, NJ: Prentice Hall, 2016.
- [19] R. G. Baraniuk. (2015). 6.341x Discrete-time signal processing. edX. [Online]. Available: https://courses.edx.org/courses/RiceX/ELEC301x/T1_2014/info
- [20] R. G. Baraniuk. (2016). RiceX: ELEC301.1x discrete time signals and systems, part 1: Time domain. edX. [Online]. Available: https://courses.edx.org/courses/RiceX/ELEC301x_/2015Q3/info
- [21] Signals and Systems. *OpenStax CNX*. [Online]. Available http://cnx.org/contents/d2CEAGW5@15.4:l-e_wSqM@24/Signal-Classifications-and-Pro.
- [22] Weather data from Jena, Germany. [Online]. Available: <http://www.bgc-jena.mpg.de/martin.heimann/weather/>
- [23] Astronomy Picture of the Day. [Online]. Available: <http://apod.nasa.gov/apod/astropix.html>



Lessons Learned from Implementing Application-Oriented Hands-On Activities for Continuous-Time Signal Processing Courses

Understanding the relationship between frequency and time domains using Fourier theory is crucial to engineers in many disciplines. In electrical and computer engineering, students usually first see this material in an introductory continuous-time signals and systems (CTSS) course and then apply it in many follow-up courses such as controls, digital signal processing (DSP), communications, and electromagnetic waves. Unfortunately, for such a crucial course, both faculty and students share longstanding concerns about the students' performance in these types of CTSS courses [1]–[3].

A collection of performance data for electrical and computer engineering students at Rose-Hulman Institute of Technology in Indiana over a span of ten years revealed that the drop/failure rate of the introductory CTSS course was several times greater than every other course in the curriculum except for the introductory electromagnetics course. Furthermore, students who did successfully complete the course received grades one to one-and-a-half letter grades lower than their average performance in other courses. We have received National Science Foundation funding to explore why it is so difficult for students to learn these concepts and to determine effective methods for helping them to grasp the concepts [4]–[9].

Digital Object Identifier 10.1109/MSP.2016.2555460
Date of publication: 1 July 2016

One possible reason for the students' learning difficulties is a lack of prior experience with the signals and systems concepts [10]. Contrast the students' experiences with a bicycle versus a cell phone. If a student is learning about mechanical advantage in the classroom, the experience of changing gears while riding a bicycle will provide some mental framework on which they can build a more theoretical foundation. However, the use of a cell phone does not have the same effect when talking about modulation because all of the energy transformations are invisible to the human senses.

We have been trying to address this lack of experience by creating application-oriented hands-on active-learning opportunities for students. There are many examples of similar opportunities described in the literature, but most of these activities make use of MATLAB, LabVIEW, or DSP hardware [11]–[16]. While these are excellent platforms, they are inherently discrete-time implementations and involve writing computer code, which creates a layer of abstraction between the experience and the theoretical concepts being studied. While we make use of these technologies, we also have designed continuous-time analog circuit platforms that allow students to

probe the underlying mechanisms of real-world applications [4], [17]. These circuits can be manipulated easily and the effects observed simultaneously in real time in both the time and frequency domains. We are not advocating the elimination of the discrete-time platforms and, in many

cases, use them for both prelab exercises and/or analysis of results. An important way for students to gain an experience with the phenomena that they are trying to model is to actually experiment with them. Using the actu-

al continuous-time systems that implement the real-world applications also creates a degree of credibility and relevance that may not be as easy to achieve with software simulations.

Just as important as the structure of the activity is the manner in which students relate the underlying theory to the phenomena that they observe during the activity. We have learned many important lessons about how to engage students and help them tie this underlying theory to the applications. For example, one problem that students have when working with applications is that they have no idea what the signals are supposed to look like so they cannot determine whether they are on the right path during the activity. Consequently, it is helpful for the prelab exercise to expose

An important way for students to gain an experience with the phenomena that they are trying to model is to actually experiment with them.

the students to the signals they will be working with during the activity.

This article describes our experiences with using these types of experiments and the lessons learned that increase their effectiveness.

Samples of application-oriented activities

When dealing with application-oriented activities, there is a tradeoff that has to be made between realism and complexity. It is difficult to teach the theory from realistic signals such as quadrature-amplitude-modulated signals or speech signals because their spectra are complicated and/or difficult to model mathematically. However, if students never see such realistic signals, they will never understand the purpose of the theory they are learning. In contrast, square waves, impulses, and other such simplistic signals are easy to model mathematically but don't always capture the variation that is seen in real-world signals. To fully understand a system, such as a cell phone or heart monitor, requires a high level of understanding of concepts such as Fourier transforms and series, system linearity and time invariance, wireless communications, electromagnetics, system frequency response, multipath, and many other concepts. To introduce such systems all at once would be overwhelming for a novice student.

When we speak of application-oriented activities, we are referring to using realistic signals operating on simplified systems that are placed into the context of these more complex systems. For example, students may explore the modulation of a speech signal, which in and of itself is not a cell phone. However, the simplified experiment provides students with the opportunity to manipulate the signals and systems in real time and observe the effects in both the frequency and time domains. Then context can be provided about how modulation of speech signals relates to how a cell phone works. It is up to the instructor to control the students' exposure to appropriate levels of detail. We have taken two different approaches to designing application-oriented activities for the students. In one approach

a single general-purpose and configurable platform, the Signals Exploration Board (SEB), was designed along with a set of weekly lab activities. In the other approach, several platforms (boards) were used to illustrate the concepts.

The SEB was designed to be easy to set up, flexible enough to relate to many types of applications, and robust enough to handle several years of student use. As shown in Figure 1, the signal path of the SEB is composed of four stages: input, sampling, filtering, and output. The input stage produces the signal $v_1(t) = v_x(t)v_y(t) + v_z(t)$, where $v_x(t)$ is one of several different input signal choices that are selected with a jumper; $v_y(t)$ can be chosen with a switch to be either 1 V, a dc voltage between -5 V and 5 V set by a potentiometer, or a time-varying signal; and $v_z(t)$ can be chosen with a switch to be either 0 V or a time-varying signal. The signal $v_1(t)$ can be sampled with various forms of sampling such as pseudo-impulse, pulse, or zero-order-hold (ZOH) sampling or the sampling can be bypassed. The filtering stage has three paths that can be selected with another jumper: a bypass with no filtering, a prewired operational amplifier for a first-order filter, or three prewired operational amplifiers for multifeedback filters that can implement up to a sixth-order filter. Students create the filter characteristics by inserting the appropriate passive devices into sockets on the board. The output stages contain a low-power driver that can drive the 50- Ω input impedance of a spectrum analyzer and a high-powered driver for 8- Ω speakers and headphones.

The SEB can be easily configured with jumpers and switches to facilitate a wide range of activities. As shown in Figure 1, the inputs header allows students to select between an applied time-varying voltage signal such as from a function generator (signal), the output of an on board microphone (MIC), their measured electrocardiography (ECG) signal, or the output of an onboard instrumentation amplifier (inst amp) that can measure differential

sensor signals such as load cells and thermocouples. All signals are relative to a common ground. These signals can then be modulated using $v_y(t)$, have noise or other signals added to them using $v_z(t)$, sampled, and filtered. The output of the system can be observed on the oscilloscope and/or spectrum analyzer and heard through a speaker or headphones. The applications that can be explored include amplitude modulation (AM) and frequency shifting, signal-to-noise ratio, harmonic distortion, music and speech signal processing, ECG signal processing, and many others.

The SEB has been used in a weekly three-hour lab during a ten-week quarter in the introductory CTSS course for the last seven years. Lab activities have been designed for each week of the quarter to give students experience with many course concepts that are more difficult to understand. The activities have gone through a lot of evolution during this period based on lessons that

were learned. In their most recent format, each activity follows a similar sequence. Students start by working with a simple signal such as an impulse train, progress to a slightly more

complicated signal such as a sum of four harmonically related sinusoids, and end with the application-oriented signal such as a speech signal.

One of the weekly lab activities begins with students exploring what happens to periodic signals as they go through a linear-time-invariant (LTI) system and ends with students measuring their own ECG signal. Students use a first-order active low-pass filter with a variety of signals. They start with a pseudo-impulse train, filter a square wave, and then measure and filter their ECG signal. In each of the stages, students are asked to complete a more open-ended activity such as to figure out how to adjust the input signal so that the filter changes the fifth harmonic by -3 dB. Students are asked to turn in a report with screen captures showing signals in both the time and frequency domains and write a paragraph that

When dealing with application-oriented activities, there is a tradeoff that has to be made between realism and complexity.

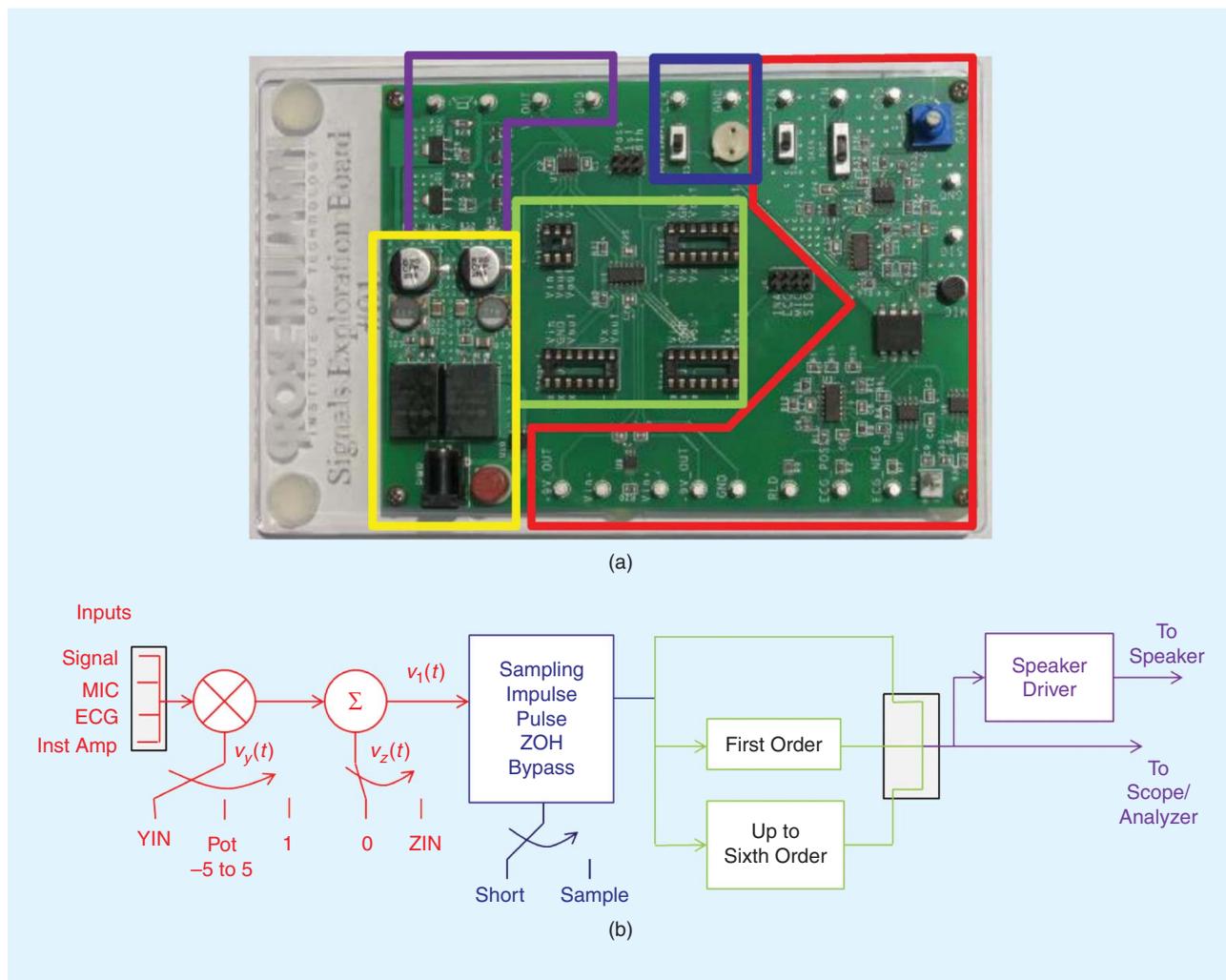


FIGURE 1. (a) A photo of the SEB with the various system parts highlighted. (b) The system diagram of the SEB. The shaded boxes with multiple inputs in the input and filtering stages are multiplexers that are implemented as headers with a single jumper that selects the signal to pass through. The arrows indicate switches that select a particular functionality.

answers problems such as “Determine and give an explanation for which filter best removed the noise from the ECG signal without altering important artifacts of the signal.” More detailed information about the SEB and the corresponding lab activities can be found in [5] and at <https://web.rose-hulman.edu/groups/SignalsEducation/SitePages/Home.aspx>.

Another approach to application-oriented activities is to explore the realistic signals by using several platforms in a controlled manner. A total of seven different activities, in addition to a MATLAB introduction, were developed for use in a 14-week semester

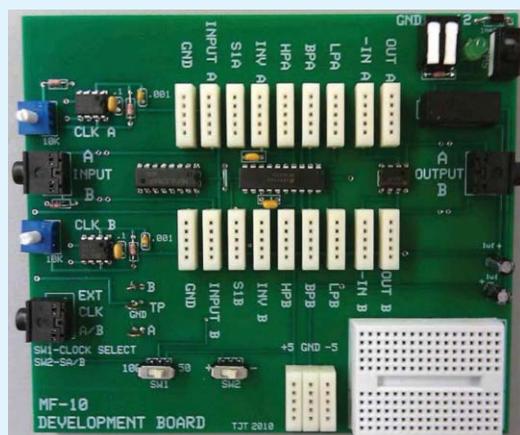
Many lessons were learned by vetting the activities with the students.

in a course with a dedicated lab component. Some applications such as satellite communications and computation of Fourier series coefficients and Fourier transforms of realistic signals were done using structured MATLAB and Simulink exercises. For convolution, the students recorded the impulse response of the campus theater using a computer

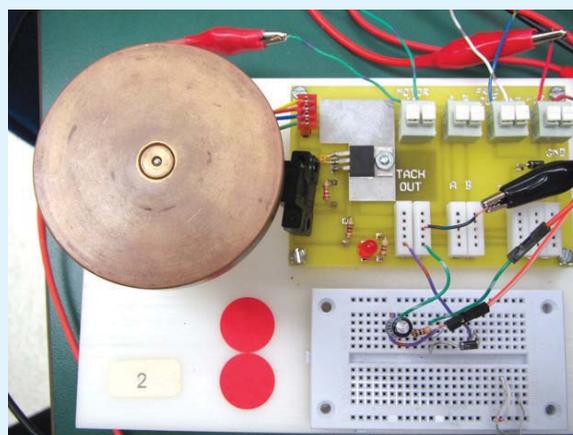
sound card and MATLAB. To study filtering, a telephone touch-tone generator and filtering platform were

developed as shown in Figure 2. The hardware platform allowed students to easily construct filters to process the two-tone signal by inserting various passive devices into the fixed headers on the board. Another hardware platform, which is also shown in Figure 2, was developed to help students better understand the Laplace transform and control systems. Students were asked to design and test a motor speed control system by modeling the motor as a first-order system. The controller and feedback were implemented using analog circuits that were constructed on the small breadboard.

Each of these activities (projects) was designed as a one-week assignment. During the activity students were given a loose set of instructions, which included



(a)



(b)

FIGURE 2. (a) A photo of hardware that allows students to easily develop analog circuit-based filters. (b) A photo of hardware that allows students to experiment with a control system.

some design components. In the touch-tone decoding and filters activity, for example, students designed a filter using the MF10 (universal monolithic dual switched-capacitor filter) chip to decode a touch-tone signal. The introduction of the activity included a discussion of the history of the development of the touch-tone telephone and related design issues. To minimize the time devoted to building the detection filters, a circuit board with an MF10 filter was given to the students. The circuit board shown in Figure 2 had the necessary clock generator, audio jacks, operational amplifiers, and an area to wire circuit components. Students were tasked with using the MF10 data sheet to create the required filter to detect a particular touch-tone. Based on their calculations, students then programmed the board by proper selection of resistors. Students were also asked to answer several short answer questions in paragraph format and analytical questions using mathematics and concepts that were presented in the course. Examples of questions for this activity include “Suppose the transmitter was nonlinear, list the first three harmonics of each frequency and then compare these harmonics with the original frequencies” and “What would happen to your detector circuit if the transmitter frequencies varied by $\pm 2\%$?”

After completing eight weeks of activities, students were asked to choose an application-oriented project based on the signals and systems concepts. During the last three weeks of the semester, the students implemented the project using hardware or MATLAB. Some project examples include active noise cancellation, a digital graphic equalizer, computing the fast Fourier transform (FFT) using a cell phone, a voice modulator, and a voice recognition control system.

Lessons learned from using application-oriented activities

Both sets of activities, described in the previous section, were initially developed in the 2009–2010 time frame but have evolved significantly since their inception. Many lessons were learned by vetting the activities with the students. The SEB has been used during three ten-week quarters per year over the last seven years and has gone through three revisions, and the activities are nothing like their original versions. The other activities using various platforms have evolved as well. The number of activities has

been reduced, and some of the activities have been extended to two weeks.

One of the more challenging aspects of working with realistic signals and systems is that most students have no idea what to expect in terms of results from their experiments. During one lab session, a student was trying to measure a speech signal but had the SEB configured incorrectly so that it was multiplying the speech signal with a low-duty-cycle pulse signal. When the instructor asked the student what was wrong with the experiment, the student responded, “I have no idea what a speech signal is supposed to look like, so how can I know what’s wrong?” If students don’t have a general sense of what

One of the more challenging aspects of working with realistic signals and systems is that most students have no idea what to expect in terms of results from their experiments.

they should expect during the experiment, there is no way they can debug the experiment. Consequently, many students end up missing the point of what they are supposed to learn from the activity. In another activity, students are asked to

measure the harmonic distortion of the SEB in the first part and then reconfigure the SEB to measure the signal-to-noise ratio. Quite often, they would submit data for the signal-to-noise ratio experiments

that came from the harmonic distortion experiments. One way to help minimize this problem is to have students do some prelab research about the realistic signals that they will be working with during the activity and then present a quiz at the beginning of the lab session. This quiz can be done in a creative way such as giving them several different signals and asking them which looks most like the signal of interest. Another method is to provide some examples of what the students should expect to see in the lab materials during the activity. This process forces them to look for details in the signals to discover the important features of the signal. Faculty members have many years of experience relating complicated signals to the mathematical theory used for examples in class, but students have no experience on which to base those relationships.

One of the primary advantages of using hardware-based activities is that they operate and can be manipulated in real time. Using the bicycle analogy, as the rider changes the gear, he/she gets an immediate sense of the change in mechanical advantage and can then associate that change with the change in gear ratio. Hardware experiments that run in real time can be manipulated easily by turning knobs or changing signals, and the effects of those changes can be observed immediately in both the time and frequency domains. Such observations can help students to more easily associate parameters such as the duty cycle of a square wave to the frequency content of the signal.

There needs to be a balance between step-by-step instructions and open-ended experimentation. While students are usually adept at taking time-domain measurements with an oscilloscope, the introductory CTSS course is usually the first time that students are taking frequency-domain measurements. Any time a new experimental technique is being used, they will need more step-by-step instruction. However, as students progress, too much detail removes the

One of the primary advantages of using hardware-based activities is that they operate and can be manipulated in real time.

necessity of making connections between the course concepts and the experience of going through the activity. For example, an initial activity may specify precisely which buttons to push on the oscilloscope to measure the FFT of a signal and at which frequencies to take measurements. Subsequent activities may simply specify which harmonics to measure or ask the students to determine at which harmonics they should measure. By leaving the experimentation more open, students must understand what a harmonic is, which harmonics are important, what features of the FFT are important, how to adjust the equipment to take the measurements, and how to interpret the results. Both of us have included design of experiments in our activities so that students are forced to apply the CTSS concepts to solve a problem.

When taking measurements with the oscilloscope, instructors should force-

fully discourage or even disable the “autoscale” function. The whole purpose of the introductory CTSS course is to make students familiar with concepts such as period, amplitude, frequency, and phase. Using the “autoscale” feature

prevents the students from thinking about what timescales and amplitudes at which they are supposed to be looking. This problem also is relevant to MATLAB when generating plots of data. Depending on the data, MATLAB can generate frequency-domain plots for signals that go far beyond any useful frequency range. For example, if looking at speech signals captured on the oscilloscope, the sampling frequency of the scope is so high that MATLAB can plot frequencies well into the hundreds of kilohertz, but the useful frequency range is below 5 kHz. Applications give students an opportunity to think about what the “useful” time-frequency-amplitude ranges are for a given problem. This is another opportunity to provide more direction about what the useful ranges are at

the beginning of the course and then let students determine these limits for themselves as the course progresses.

Dealing with application-oriented experiments means that there are many opportunities for the experiment to go wrong or become more difficult. There are many times when students will see phenomena that are not easy to explain with their current level of knowledge or are excellent examples of the topics covered in class but are difficult for students to connect with the theory. These are excellent opportunities to show the whole class at one time and give them a chance to try to explain the phenomena being observed. One common example is dealing with triggering of low duty-cycle signals or signals with discontinuities such as ramps and pulses. Another example is if the system becomes dysfunctional and the signals are different than the expected result. The class as a whole can be asked to figure out what aspects of the signal indicate the dysfunction and if they can determine the dysfunction from the signal.

Yet another common example is when periodic signals are not exactly periodic such as when measuring an ECG signal or recording the sound of a musical instrument playing a single note.

While participating in these application-oriented activities, students observe or experience phenomenon and then must learn to use the language introduced in the CTSS course to explain that phenomenon. In many ways, this process of observation and explanation is similar to learning a foreign language, and the best way to learn a new language is to use it.

Writing to explain observations

We have made extensive use of short-answer questions for which students must explain their observations in paragraph form in addition to analytical mathematics. There is literature that supports the use of short-answer questions to both improve and assess the students’ advanced mathematical reasoning skills [18]. An example question that is asked during the modulation activity is, “Both modulation and time-scaling make a system time-varying and can alter the frequencies of the input

signal. Describe how they sound different and relate these differences to any changes in the fundamental frequency of the output signal." Students generally are very poor at using the correct terminology for explaining the phenomenon that they observe and often apply the incorrect concepts to an observation. The process of explaining their observations with words helps to check their own understanding of the concepts and their observations. To improve their understanding, the students must receive formative assessment of their answers. These essay questions can be added to examinations to encourage them to put in the effort to ensure that their understanding is correct. Because of the extra effort involved with the short-answer questions, it is helpful to focus on just one or two per activity that capture the key concepts. In addition, we have asked students to prepare a short oral presentation to the class to explain the basic concepts and their findings.

Conclusions

For the past eight years, we have used application-oriented activities in the introductory CTSS course. Such activities can help students to connect the mathematical theory learned in the course to how the theory is applied in real-world applications. The level of detail in the activity needs to be carefully monitored by the instructor to not overwhelm the students and make it easier for the students to connect the theory to the application. Using hardware-based activities can make it easier for students to associate the theory to the application by facilitating real-time cause-effect relationships. The manner in which the application-oriented activities are performed can have a tremendous impact in how well students are able to better understand the theory.

Acknowledgments

This work is sponsored by National Science Foundation grant 1140995. As another part of this grant we conducted student interview studies, and the analysis of those studies led to many of the insights presented in this article. We would like to thank Dr. Ruth Streveler and Dr. Farrah Fayyaz from the Purdue School of Engineering Education who developed, conducted, and analyzed the results of those studies.

Authors

Mario Simoni (simoni@rose-hulman.edu) received his Ph.D. degree from the Georgia Institute of Technology. He is the department chair and professor of electrical and computer engineering at Rose-Hulman Institute of Technology in Terre Haute, Indiana, where he has been teaching since 2001.

Maurice Aburdene (maurice.aburdene@bucknell.edu) received his Ph.D. degree from the University of Connecticut. He is a professor of electrical and computer engineering at Bucknell University, Lewisburg, Pennsylvania, where he has been since 1981.

References

- [1] J. Nelson, M. Hjalmarson, K. Wage, and J. Buck, "Students' interpretation of the importance and difficulty of concepts in signals and systems," in *Proc. 2010 IEEE Frontiers in Education Conf. (FIE)*, Oct. 2010, pp. T3G-1–T3G-6.
- [2] K. Wage, J. Buck, and C. Wright, "Obstacles in signals and systems conceptual learning," in *2004 IEEE Digital Signal Processing Workshop/3rd IEEE Signal Processing Education Workshop*, Aug. 11, 2004, pp. 58–62.
- [3] M. Simoni, F. Fayyaz, and R. Streveler. (2014, June). Data mining to help determine sources of difficulty in an introductory continuous-time signals and systems course. presented at the 2014 ASEE Annu. Conf. [Online]. Indianapolis, IN. Available: <https://peer.asee.org/20244>
- [4] M. Simoni, "Work in progress—A hardware platform for a continuous-time signals and systems course," in *Proc. 2011 IEEE Frontiers in Education Conf. (FIE)*, Oct. 12–15, 2011, pp. S4G-1–S4G-2.
- [5] M. Simoni, M. F. Aburdene, and F. Fayyaz. (2013, June). Analog-circuit-based activities to improve introductory continuous-time signals and systems courses, presented at 2013 ASEE Annu. Conf.

[Online]. Atlanta, GA. Available: <https://peer.asee.org/19202>

- [6] F. Fayyaz, "A qualitative study of undergraduate electrical engineering students' problematic reasonings in 'continuous time signals and systems' courses," Ph.D. dissertation, School of Engineering Education, Purdue Univ., West Lafayette, IN, 2015.
- [7] M. Simoni, M. F. Aburdene, F. Fayyaz, V. A. Labay, J. Wierer, and W. Huang, "Improving learning in continuous-time signals and systems courses through collaborative workshops," presented at 2015 ASEE Annu. Conf. Exposition, Seattle, Washington, June, 2015.
- [8] M. Simoni, M. Aburdene, and F. Fayyaz, "Why are continuous-time signals and systems courses so difficult? How can we make them more accessible?" in *Proc. 2013 IEEE Frontiers in Education Conf.*, Oct. 23–26, 2013, pp. 6–8.
- [9] M. Simoni, M. Aburdene, and F. Fayyaz, "Workshop: Understanding learning difficulties in continuous-time signals and systems courses and making these courses more accessible," in *Proc. 2013 IEEE Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)*, Aug. 11–14, 2013, pp. 308–310.
- [10] D. A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*, Englewood Cliffs, NJ: Prentice Hall, 1984, vol. 1.
- [11] L. G. Huettel, "A DSP hardware-based laboratory for signals and systems," in *IEEE 12th Digital Signal Processing Workshop and 4th Signal Processing Education Workshop*, Moose, WY, 2006, pp. 456–459.
- [12] J. E. Corter, J. V. Nickerson, S. K. Esche, C. Chassapis, S. Im, and J. Ma, "Constructing reality: A study of remote, hands-on, and simulated laboratories," *ACM Trans. Comput.-Hum. Interac.*, vol. 14, no. 2, pp. 711–727, Aug. 2007.
- [13] G. Ferre, A. Giremus, and E. Grivel, "Small-group learning projects to make signal processing more appealing: From speech processing to OFDMA synchronization," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2009, pp. 2317–2320.
- [14] C. H. G. Wright, M. G. Morrow, M. C. Allie, and T. B. Welch, "Enhancing engineering education and outreach using real-time DSP," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2008, pp. 2657–2660.
- [15] J. Turner and J. P. Hoffbeck, "Putting theory into practice with Simulink," in *Proc. American Society for Engineering Education Annu. Conf. Exposition*, Portland, OR, 2005.
- [16] N. A. Pendergrass, "Using computers, simulators and sound to give hands-on experience," in *Proc. American Society for Engineering Education Annu. Conf. Exposition*, Washington, D.C., 1996.
- [17] M. F. Aburdene and K. Nepal, "Wow! Linear systems and signal processing is fun!" in *Proc. 2011 Frontiers in Education Conf. (FIE)*, Oct. 12–15, 2011, pp. F1G-1–F1G-6.
- [18] T. Dreyfus, "Advanced mathematical thinking processes," in *Advanced Mathematical Thinking*, D. Tall, Ed, Dordrecht: Kluwer Academic, 1991, pp. 25–41.



Understanding and Predicting Epilepsy

Epilepsy is the second-most prevalent neurological disorder after migraine and affects 1–2% of the world population. Epilepsy is a spectrum of disorders defined by the occurrence of epileptic seizures, which are characterized by the abnormal firing of large populations of neurons. Epilepsies can have a genetic component (as a direct cause or as risk factor) or can be triggered by a brain insult (including stroke, meningitis, and brain trauma). After the initial insult, complex reorganizations occur in neuronal networks (epileptogenesis), ultimately favoring the emergence of spontaneous seizures [1]. Epilepsy is a dynamic process as the reorganization continues during one's life. Other brain regions may get involved, possibly invoking different mechanisms, requiring changes in medication. Seizures can be very difficult to control, and 30% of patients are drug resistant. Epilepsy research thus faces two major challenges:

- *Understanding the basic mechanisms underlying seizure genesis.* This is a key issue if one wants to design new therapeutic solutions to prevent the occurrence of seizures (in particular to treat drug-resistant patients) and if one wants to identify which brain regions produce seizures in a given patient (in particular to plan neurosurgery).
- *Predicting seizures.* If seizure control cannot be achieved, it is equally crucial to warn patients of incoming seizures. Patients could inform people

around them and place themselves in safe conditions.

Despite decades of research, we do not really know how seizures start, propagate, and terminate, and we still do not understand how a “normal” brain becomes “epileptic.” Using human tissue and experimental animal models, we learned that the molecular architecture of epileptogenic networks is considerably modified. Hundreds to thousands of proteins are up- or down-regulated in epilepsy. But this reorganization is so complex that we don't know how to interpret such an amount of data, i.e., it is difficult to determine which of these modifications are causally related to seizure genesis and propagation. One way to make sense of it is to use a computational approach to identify key parameters. As will be developed next, such a computational approach is not straightforward, and some guiding principles should be proposed.

Considerable efforts have been made to address the second challenge, predicting seizures. Numerous approaches/algorithms have been developed, but so far, none is successful enough for clinical use.

Why are breakthroughs so difficult to obtain in epilepsy research? Two main challenges need to be considered:

- 1) There are multiple possible mechanisms underlying seizure genesis and propagation.
- 2) We overrely on one type of observation (electrophysiological recordings) to build our conceptual frameworks or theories.

Historically, we have used electrophysiological signals as a gold standard, i.e., as objective markers of seizures. Electrophysiological signals are time-dependent

fluctuations in field potential due to the movement of charged particles in the neuronal tissue. These fluctuations are a highly integrated signal as compared to the biological processes that gave rise to them. Arguably, the answers to our questions need

Epilepsy is a spectrum of disorders defined by the occurrence of epileptic seizures, which are characterized by the abnormal firing of large populations of neurons.

to be investigated at the molecular scale, since brain activity is molecular by nature. To understand the nature of the problem, let's consider the classical example of the fluttering of a butterfly's wings that unleashes a storm. If weather scientists only have access to air temperature and pressure data, they may never be able to identify the cause of the storm. The analogy holds for seizures. The brain is a highly interconnected complex molecular system. Electrophysiological signals may hold some clues (i.e., key molecular events may have an electrophysiological trace), but, so far, we don't really know what to look for.

Future progress may require accepting the multiplicity of solutions and focusing on molecular events. Theoretical studies clearly showed the complexity of the problem and already provided important insights.

Modeling seizures

Two broad types of modeling approaches are commonly used: detailed and lumped models [2]. Microscopic detailed models try to be as realistic as possible, including a maximum of biophysical details. Neurons can be modeled as multicompartment structures, including a large variety of ionic and ionotropic channels with specific spatial distributions along the somato-dendritic tree. Roger Traub pioneered this approach, and as computational power increased, it has been possible to include an increasing number of neurons and parameters. Ivan Soltesz's laboratory was able to model a whole part of the hippocampus with hundreds of thousands of neurons, using exquisite detailed network architecture. The caveat of this approach is the size of the parameter space. Many parameters have never been measured (e.g., the deactivation curve of ion channel X in the distal dendrite of interneuron type Y). They must be guessed from other values measured in other types of cells. Even if a given parameter has been determined, we must choose from a distribution of values measured experimentally, characterized by a mean and a standard deviation (measures of the same parameter in different experiments invariably lead to different results). Which value should be used, the mean, the median, or the extremes?

When trying to understand how the simple rhythms could emerge in the stomatogastric ganglion network, Eve Marder faced the same challenge. The approach she took was to explore the whole parameter space. Each parameter could take a finite number of values, and she extracted all sets of parameters, which could produce the same rhythm in silico as was observed in vivo. She demonstrated that there exist a huge number of parameter configurations, all giving rise to in vivo-like network activity [3]. The same concept can be extended to seizures in the temporal lobe. Although

extremely time consuming, it would be particularly interesting to use “à-la-Marder” strategy, and determine which sets of parameters can give rise to seizure-like events (e.g., in a detailed model of the hippocampal network), and compare these sets to those giving rise to physiological rhythms (e.g., theta or gamma oscillations) in the same model. Perhaps certain parameters are functional “hubs,” i.e., their modification would consistently lead to seizure genesis/propagation regardless of alterations in other parameters. Such predictions could then be tested experimentally. There is, however, a major difficulty to solve. In Marder's work, the activity in the stomatogastric network can easily be described and quantified (she used some well-established metrics: interburst frequency, intraburst frequency, refractory period, etc.). Except for absence seizures, seizures with focal onset do not appear to follow general rules of organization. Hence, how can we accept as a seizure-like event, the activity generated in silico? Pure mathematical approaches may give us some clues.

Lumped (e.g., neural mass/field) models are based on the assumption that temporal and spatial averages are sufficient to characterize the dynamics of neuronal networks. They are more focused on general rules than on biophysical details. Hence, their predictive value is limited to general principles, since by nature they manage to reduce the size of the parameter space. However, the general rules thus obtained can be used as guiding principles for detailed models, as developed hereafter.

Most modeling approaches tried to reproduce seizure genesis/propagation by appropriate modifications of model parameters. Few attempted to understand the nature of seizure genesis/propagation. Clinicians have described many forms of epilepsies, and they have underlined the difficulty that they may encounter to achieve seizure control in patients. Hence, it is generally assumed that seizures are very complex phenomena. But are they?

Two important clinical observations can give us some hint about the way we should phrase the problem:

- Seizures are found in numerous neurological disorders, e.g., Alzheimer's disease, Huntington's disease, and autism. Why are seizures so common in other diseases?
- Any normal brain can be forced to have a seizure, for example, after an electroconvulsive shock. Hence, a seizure is a type of physiological activity; it is hardwired in neuronal networks.

Seizures are hardwired physiological activities

Epileptic seizures can be triggered and recorded across species (from flies to humans) and brain regions. Interestingly, seizures with focal onset share similar properties: the presence of fast oscillations and the occurrence of slower spike-and-wave discharges (Figure 1). Each type of activity occurs with a different time scale (frequency), fast and slow, respectively. The activity of neuronal networks can be represented by a time series expansion of the electrophysiological signals and described by state variables. Two state variables are

sufficient to describe fast oscillations, and two other state variables are sufficient to describe slower spike-and-wave discharges. Since, by definition, seizures are a recurring phenomenon in epilepsy, it is possible to introduce a fifth state variable evolving on a very slow time scale, driving neuronal networks to seizure onset, controlling the seizure dynamics, and its offset. Such a fifth variable allows the system to switch autonomously from control to seizure activity and from seizure activity back to control activity. Five state variables are thus sufficient to account for the properties of seizures with partial onset (i.e., dynamics and constituents in terms of fast oscillations and spike and wave discharges). The resulting model is called the *epileptor* [4], and its equations are shown in Figure 1.

It is generally assumed that seizures are very complex phenomena. But are they?

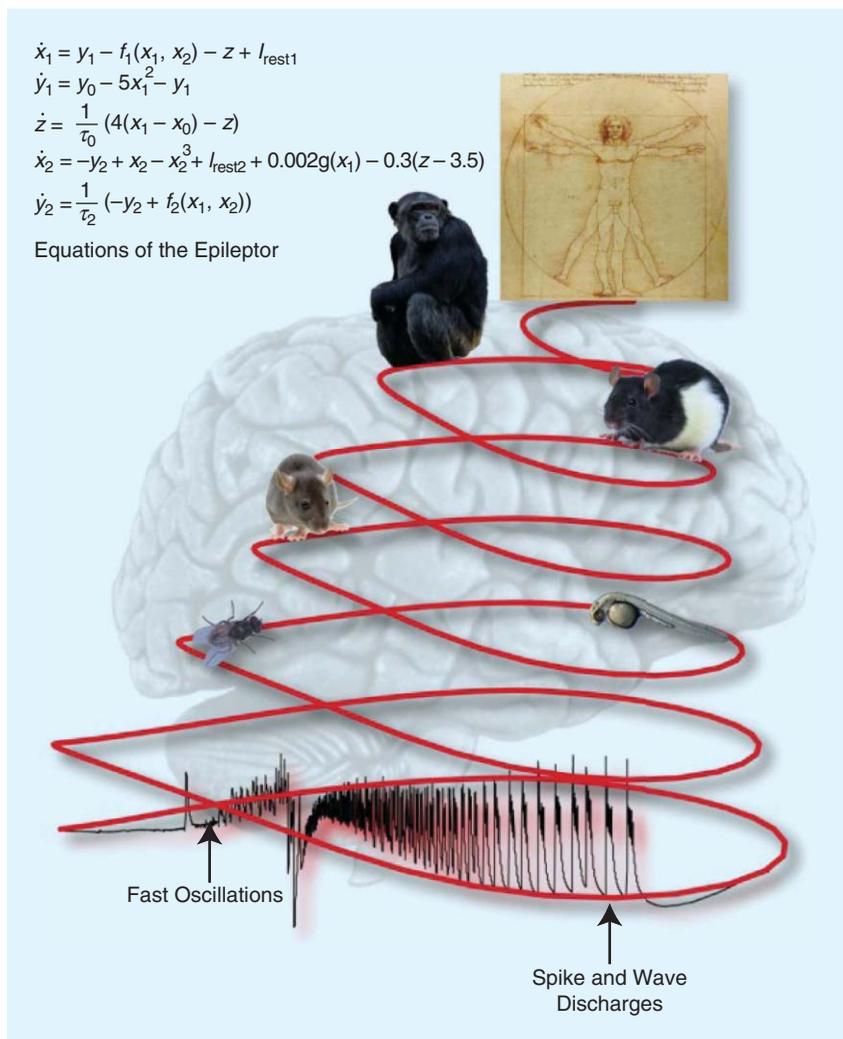


FIGURE 1. The principles of the dynamics of seizures with a focal onset. The bottom trace is the recording of a seizure in the hippocampus of a mouse. Two patterns are clearly apparent: fast discharges (or oscillations) and spike and wave discharges. The top left panel shows the equations of the epileptor. The two state variables x_1 and y_1 describe the fast oscillations, and x_2 and y_2 describe the spike and wave discharges. The z state variable leads the system to seizure onset and drives the seizure dynamics until its offset. As written, the equations predict a saddle node bifurcation at seizure onset and a homoclinic bifurcation at seizure offset. These predictions were verified in various species and brain regions, including in patients with different types of seizures with focal onset. Geometrically, seizures are spirals traveling on a cone (when projected in three dimensions).

spiral on a cone (Figure 1), a very simple geometrical object.

Based on the work of Izhikevich, we predict that there are at least 16 types of seizures. Although 83% of seizures recorded in drug-resistant patients belong to the saddle node-homoclinic class, it will be important to identify which other classes are also represented. This is particularly essential to consider in terms of mechanistic insight. The epileptor is a phenomenological model, i.e., it does not claim any biophysical relevance. However, it imposes strong constraints on the behavior of the network at seizure onset and offset. As mentioned previously, there is no consensual definition of what a seizure should look like (in contrast to basic brain rhythms such as theta and gamma oscillations). In the absence of a well-defined metric, it is difficult to assess whether seizures obtained *in silico* bear any physiological relevance. The epileptor shows that a “dynamic” metric in terms of bifurcations can be considered. When using a very detailed model (or even a lumped model) to study a specific type of seizure, the set of parameters must satisfy the properties of the bifurcations (e.g., logarithmic slowing down of the activity for a homoclinic bifurcation at seizure offset for the main class of seizures measured in patients). This should greatly limit the size of the parameter space.

The epileptor also provides a different approach to tackle seizure mechanisms. Since seizure onset occurs via a bifurcation, it means that the trajectories of brain activities need to cross a certain threshold (or barrier of energy). It can be argued that, to understand how seizures start, it is sufficient to identify the forces that drive the network over the threshold. It is equally important to determine the forces that drive the network back to a “normal” state at seizure offset. If such forces can be identified, specific interventions may be designed to prevent reaching seizure onset. Likewise, protocols could be designed to abort seizures as soon as they start, as successfully demonstrated with optogenetic approaches by the groups of John Huguenard, Dimitri Kullmann, and

In the model, seizure onset and offset occur via bifurcations. When looking at the mechanisms of bursting, Eugene Izhikevich identified four possible bifurcations at bursting onset and four at bursting offset. There are thus 16 possible types of bursts. Generalizing this concept to seizures, the epileptor equations correspond to one class of these 16, with saddle-node and homoclinic bifurcations at seizure onset and offset, respectively. These bifurcations were found across brain regions and

nonhuman species and in 83% of drug-resistant patients [4]. Interestingly, seizures triggered in “healthy” brains were characterized by the same bifurcations, supporting the universal nature of the model and the fact that seizures are endogenous brain activities. Topologically, when projected in three dimensions (i.e., using the first state variable describing fast oscillations, the first state variable describing the spike and wave discharge, and the fifth very slow variable), a seizure with partial onset is a

Ivan Soltesz. Finally, the other key concept is that of the threshold. It is important to assess the threshold of a given's individual (to assess their susceptibility to seizures). It is also essential to measure at which distance the brain trajectories are from the threshold when trying to devise ways to predict seizures. Such knowledge regarding driving forces and thresholds is proving very challenging to obtain. The difficulty stems from the multiplicity of possible forces and thresholds, which may be constitute the core reason why seizures may be so difficult to treat.

All roads lead to seizures

The key concept is that there are multiple different ways to cross multiple thresholds, always ending up with the same type of seizure (Figure 2). For example, overdoses or electroconvulsive shocks are external forces that can trigger similar forms of seizures in humans, yet their underlying biophysical mechanisms are totally different. In addition, the sensitivity to these forces varies from one individual to the next, meaning that the thresholds related to these forces are specific to a given individual. We provided direct experimental evidence of the existence of multiple thresholds in a neuronal network [4]. Using the same hippocampus, we first triggered a seizure-like event by increasing the level of synaptic noise in the preparation. When the noise reached a critical value (stochastic resonance) a seizure-like event occurred. Returning to baseline conditions, we then increased the osmolarity, thereby triggering a second seizure-like event similar to the first one. In that case, there was no change in synaptic noise, as a different mechanism was involved in seizure crossing. Finally, when we combined individual sub-threshold conditions for noise and osmolarity, we triggered a seizure-like event via a third pathway. These results demonstrate that multiple conditions (hence multiple mechanisms) can lead to the same seizures in the same preparation [4]. Despite the multiplicity of possibilities, the dynamics of seizures remains invariant.

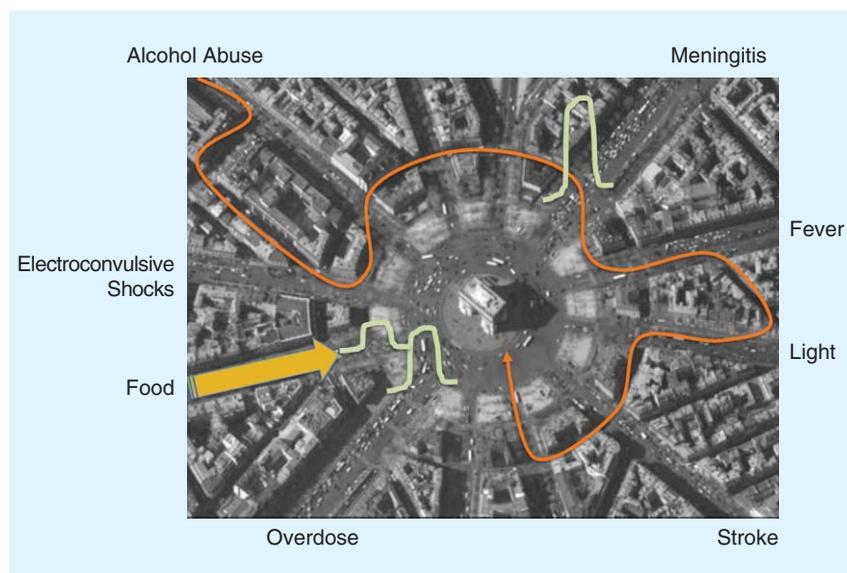


FIGURE 2. Different paths/mechanisms can lead to the same type of seizure. Seizures are part of the landscape of possible brain activities; they are hardwired in neuronal networks from flies to humans. Multiple roads can lead to the same end point (the seizure). In humans, seizures can be triggered by alcohol abuse, meningitis, fever, light (a reflex epilepsy), stroke, overdose, food poisoning, and electroconvulsive shocks. Each path is characterized by a specific threshold. In the case of an established cause, e.g., food poisoning with domoic acid in mussels, domoic acid would act as a “force” driving the network above seizure threshold. Each path is also characterized by specific underlying mechanisms, e.g., domoic acid and fever are likely to act via very distinct mechanisms. When there is no obvious causal event, which is usually the case when seizures occur spontaneously, the trajectory may be very complex (orange line). From one seizure to the next, the trajectory is not necessarily the same. In that case, the underlying mechanisms would be different. Perhaps the complexity to treat epilepsy stems from such multiplicity of entry points and mechanisms.

Therefore, there are multiple entry points to the seizure state, with each characterized by its own threshold and, hence, its own underlying biophysical mechanisms. In a given patient, it is possible that different thresholds are weakened, thus providing many different seizure entry points. We can make the reasonable assumption that the modus operandi of antiepileptic drugs is to increase the threshold and/or to act negatively on the forces that drive the networks close to a threshold. Since current antiepileptic drugs cannot control all entry points, we can speculate that drug-resistance is due to the fact that there are too many possible entry points in some patients.

Theoretical models could play a determinant role here. Instead of using hypothesis-driven approaches (an imbalance

between excitation and inhibition, a downregulation of a given ion channel, etc.), it would be important to run a systematic search for seizure thresholds (their nature, their location) in the network. Since we are dealing with a complex interacting system at the molecular level, such thresholds may be looked for in the main pathways that can have major

Despite its caveats, signal processing can bring some deep insight into seizure mechanisms.

and widespread influence on network activity, such as metabolism, inflammatory response, ion homeostasis, and hormones. Such predictions may then be

tested in experimental models (the data may already exist in the literature) or even in patients, which could lead to new therapeutical approaches. Taking a step in this direction complements the traditional electrophysiological approach with a molecular one. Yet, electrophysiology provides invaluable information. Many

methods have been designed to extract the maximum of information from it.

Analyzing electrophysiological signals

In the clinic, electroencephalography is central for the diagnosis of epilepsy. It is essential when neurosurgery is the last remaining hope for patients with drug-resistant epilepsy. Neurosurgery is considered successful if patients become seizure-free or at least controlled with antiepileptic drugs. Before attempting neurosurgery, it is necessary to precisely delineate the epileptogenic zone (EZ), i.e., the set of brain regions that is responsible for seizure genesis. Seizures are a network phenomenon, and the EZ is often made of multiple different brain regions. Seizures can thus originate from various regions in a given patient and then propagate to another set of regions (the propagation network). The difficulty is to determine the exact extent of the

EZ. In such a clinical context, this means unraveling seizure mechanisms in a patient-specific manner. If the EZ is not properly assessed, neurosurgery may fail, which happens in 30% of the cases on average. Since neurosurgery consists of removing parts of the brain, it is essential to remove only what is necessary, and thus limit possible subsequent functional deficits. Many patients cannot be operated on because their EZ includes eloquent cortex (like motor and language cortex). Mapping the EZ can be a very difficult endeavor (as assessed by the failure rate). Clinicians use a number of modalities, including imaging methods, but the gold standard remains intracranial electrophysiological (iEEG) recordings. This technique is the only one providing direct access to human brain activity in situ. Since it is not possible to place electrodes in every possible region, a number of sites are chosen before electrode implantation based on a first clinical assessment using noninvasive modalities (surface

Experimental evidence demonstrates the multiplicity of solutions, supporting the notion that the way networks approach seizure thresholds is not universal.

EEG, imaging). Patients are recorded during several days/weeks to measure a maximum of spontaneous seizures. A visual inspection of electrophysiological signals can give some hints of the possible size of the EZ. However, in many instances, it is difficult to determine whether a region belongs to the EZ or to the propagation zone (it is assumed that the propagation zone does not include regions from where the spontaneous seizures can emerge—seizures just pass through these regions).

Many computer-based analysis techniques have been developed to identify the EZ based on multisite iEEG recordings [5]. These techniques make use of various methods developed in the field of

information processing. Usually, these methods try to unravel the relationships between the different signals (i.e., between the different recorded regions). For example, cross-correlation measures the similarity of two time series as a function of a

time-lag. Coherence identifies significant frequency-domain correlation between the two time-series. Mutual information quantities show the shared information between two time series. The h2 method calculates nonlinear correlation coefficients. Transfer entropy measures the connectivity if the information of one time series can reduce the degree of uncertainty about future values of another.

However, there is no ideal method. Using computer-generated signals, for which the interdependency between the signals was known, a systematic test of 42 different methods demonstrated that all methods failed to correctly identify the relationships between signals [6]. In addition, all methods need the setting of some parameters (time lag, frequency band, etc.), which values cannot be known a priori [6]. This may explain why four different analysis algorithms developed by four independent groups gave very different results regarding the evaluation of EZ when using the same clinical data set [5].

Hence, despite the amount of efforts put in designing complex signal processing methods, there is no good solution available to clinicians to evaluate the EZ. Perhaps the main difficulty lies in the fact that there is no available ground truth to validate any of these methods. The validation either comes from the visual inspection by the clinicians and/or indirectly because the removal of brain areas made the patient seizure-free. There is no known objective procedure to characterize the EZ. So far, the most reliable readout would be a decrease in surgery failure rate based on the predictions obtained from computer models.

Despite its caveats, signal processing can bring some deep insight into seizure mechanisms. However, the traditional spatiotemporal approach is favorable with which to start. Here we define spatial as making use of all recording sites (including those performed in assumed “healthy” regions) and temporal as taking into account the dynamics of signals far from, close to, during, and after a seizure. Since we don’t know where and when the markers of the EZ are to be found, it is important to consider all possible options. The same principles apply to seizure prediction.

Seizure prediction—Which signal to process?

Since seizure control cannot be achieved in 30% of patients, it would be very important to be able to predict incoming seizures, if only to provide patients with a warning signal. Seizures can be life threatening, and a warning signal would enable patients to inform people around them or give them enough time to ensure their own safety. It is assumed that EEG signals contain some type of hidden information about incoming seizures. That is to say, with appropriate signal processing, one would be able to extract from the EEG the necessary predictive markers. After two decades of extensive work, and multiple forms of signal processing applied to EEGs, the results are rather negative [7]. Computer models can be trained on specific data sets, but they

fail when applied to different sets of patients [7]. One alternative solution is to train the algorithm on a specific patient during a certain time and use this information to predict the incoming seizures in the same patient (personalized medicine). This strategy was used on a set of patients equipped with electrocorticography grids for long-term recordings and a wireless system to transfer EEG signals to a data processing unit [8]. After training the system during several weeks, the incoming seizures could be reasonably and reliably detected provided that the detection threshold was maintained at a low value. But the major caveat was a high rate of false alarms [8]. Reliable seizure detection would open the way to a closed-loop system to stop seizures, e.g., with neurostimulation [9].

Perhaps the analysis of electrophysiological signals is not the best way to predict seizures and to understand the mechanisms underlying their genesis. The fact that clinicians and basic researchers focus on EEG signals has a historical origin, with the discovery that brain activity could be characterized by electrical signals. As mentioned previously, EEG signals reflect a highly integrated flux of charged particles, mostly due to synaptic activity. Since a seizure can be objectively characterized at the EEG level, it was assumed that some changes in network activity would occur before the seizure. As mentioned in this article, it is possible to postulate the existence of a “force” driving neuronal networks toward seizure threshold [4]. The main characteristic of this force is to evolve on a very slow time scale (the fifth state variable in the epileptor), which naturally points at slow molecular processes. The biophysical correlates of this

If the paths leading to seizures are very diverse in a given patient, their fingerprints may also be different.

slow variable remain unknown. But some have already been identified during seizures, including extracellular K^+ and O_2 concentrations and molecules linked to energy metabolism, like adenosine triphosphate production [4]. It would be particularly interesting to monitor such molecular activities in vivo before, during, and after seizures.

Since seizure crossing can occur at multiple locations, the “force,” and hence its biophysical mechanisms, may vary from one seizure to the other, evolve in time as a function of environmental factors, or as the nature of the epilepsy changes during the patient’s lifetime. These arguments may also explain why seizure prediction based on EEG signals has mostly failed, as its hidden assumption is that preictal states follow rules that are universal across patients and seizure types. Experimental evidence demonstrates the multiplicity of solutions, supporting the notion that the way networks approach seizure thresholds is not universal [4].

Conclusions

Epilepsy research has been conducted with the firm belief that magic bullets may be found to treat patients (with the ultimate drug), predict seizures (the key EEG biomarker), and identify the epileptogenic zone (the key algorithm). Perhaps, it is time to accept the complexity and multiplicity of solutions. Marder’s work is instrumental in that respect. Her laboratory rigorously demonstrated that there exists a huge number of network configurations (or detailed molecular architectures) giving rise to exactly the same type of network activity [3]. Seizures being an activity endogenous to most neuronal networks, they are multiple ways to produce them. If the paths leading to seizures

are very diverse in a given patient, their fingerprints may also be different. Perhaps the solution lies in the use of multimodal approaches, monitoring different paths/mechanisms simultaneously. This would require the development of new technological tools [10], [11] and conceptual approaches.

Author

Christophe Bernard (christophe.bernard@univ-amu.fr) is the director of research at Institut de Neuroscience des Systèmes, Inserm UMR_S 1106, Aix Marseille Université, France.

References

- [1] E. M. Goldberg and D. A. Coulter, “Seizing the opportunity: Stem cells take on epilepsy,” *Cell Stem Cell*, vol. 15, pp. 527–528, Nov. 2014.
- [2] C. Bernard, S. Naze, T. Proix, and V. K. Jirsa, “Modern concepts of seizure modeling,” *Int. Rev. Neurobiol.*, vol. 114, pp. 121–153, Aug. 2014.
- [3] A. A. Prinz, D. Bucher, and E. Marder, “Similar network activity from disparate circuit parameters,” *Nat. Neurosci.*, vol. 7, no. 12, pp. 1345–1352, Dec. 2004.
- [4] V. K. Jirsa, W. C. Stacey, P. P. Quilichini, A. I. Ivanov, and C. Bernard, “On the nature of seizure dynamics,” *Brain*, vol. 137, pp. 2210–2230, Aug. 2014.
- [5] R. G. Andrzejak, O. David, V. Gnatkovsky, F. Wendling, F. Bartolomei, S. Francione, P. Kahane, K. Schindler, and M. de Curtis, “Localization of epileptogenic zone on pre-surgical intracranial EEG recordings: Toward a validation of quantitative signal analysis approaches,” *Brain Topogr.*, vol. 28, no. 6, pp. 832–837, Nov. 2015.
- [6] H. E. Wang, C. G. Benar, P. P. Quilichini, K. J. Friston, V. K. Jirsa, and C. Bernard, “A systematic framework for functional connectivity measures,” *Front. Neurosci.*, vol. 8, pp. 405, Dec. 2014.
- [7] W. Stacey, M. Le Van Quyen, F. Mormann, and A. Schulze-Bonhage, “What is the present-day EEG evidence for a preictal state?” *Epilepsy Res.*, vol. 97, no. 3, pp. 243–251, Dec. 2011.
- [8] M. J. Cook, T. J. O’Brien, S. F. Berkovic, M. Murphy, A. Morokoff, G. Fabinyi, et al. “Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: A first-in-man study,” *Lancet Neurol.*, vol. 12, no. 6, pp. 563–571, June 2013.
- [9] M. Morrell, “Brain stimulation for epilepsy: Can scheduled or responsive neurostimulation stop seizures?” *Curr. Opin. Neurol.*, vol. 19, no. 2, pp. 164–168, Apr. 2006.
- [10] A. Williamson, J. Rivnay, L. Kergoat, A. Jonsson, S. Inal, I. Uguz, et al. “Controlling epileptiform activity with organic electronic ion pumps,” *Adv. Mater.*, vol. 27, no. 20, pp. 3138–3144, May 2015.
- [11] D. Khodagholy, T. Doublet, P. Quilichini, M. Gurfinkel, P. Leleux, A. Ghestem, et al. “In vivo recordings of brain activity using organic transistors,” *Nat. Commun.*, vol. 4, 1575, 2013.

SP

TIPS & TRICKS

Zhe Chen, Shuwen Wang, and Fuliang Yin

A General Design Method for FIR Compensation Filters in Δ - Σ ADCs

In this column, we describe a new design method for a precise finite impulse response (FIR) compensation filter as a component in a delta-sigma modulator (DSM), also referred to as a Δ - Σ modulator. Several cost functions are first proposed based on their compensation purpose. Then the steepest descent method is adopted to find out the optimum filter coefficients. By applying this method, it is possible to achieve a relatively flat frequency response magnitude in the passband even if the order of the filter is low.

A Δ - Σ modulator-based analog-to-digital converter (ADC) integrates sampling, quantization, and coding through digital signal processing. By relying on oversampling and quantization noise shaping, the in-band noise is pushed to the high frequency range, where it is removed by a proper prefilter and downsampler [1]. A Δ - Σ ADC presents multiple advantages, such as high precision (24-bit resolution or more), linearity, robustness to integrated circuit (IC) chip manufacturing errors, etc. Therefore, it has become a mainstream technique of modern precise ADCs, where it has been used in a wide range of application areas, including communications, digital audio systems, multimedia systems, radar, and sonar.

The oversampling factor (the ratio of the Δ - Σ modulator's output sampling frequency to that of the ADC) is usually big to obtain high resolution

in a Δ - Σ ADC (e.g., the ratio may be more than 2,000 in the AD7705 chip manufactured by Analog Devices, Inc.). To reduce the computational complexity and memory usage [2], [3], usually a group of prefilters and downsamplers are sequentially and alternatively applied to the output signal of a Δ - Σ modulator, as shown in Figure 1(a), where LPF_n and $\downarrow D_n$ denote the lowpass prefilter and downsampler with a factor D_n in the n th stage, respectively. A cascaded-integrator-comb (CIC) filter usually appears as a lowpass filter in the first stage of a Δ - Σ ADC due to its simplicity, which requires no multiplication or coefficient storage [4], as shown in Figure 1(b). Unfortunately, the frequency response magnitude (FRM) of CIC filters is not flat enough for most applications. Therefore, a compensation filter

is needed to flatten the FRM of a CIC filter, as shown in Figure 1(c).

Several methods have been presented for CIC compensation filter design. The compensation methods based on the sharpening technique [5] and the interpolated second-order polynomials [6] are the most classical ones, reducing the maximum ripple of passband down to 0.004 dB. The nonuniform frequency sampling method based on the least-square technique was adopted in [4], in which the FIR compensation filter offsets the ripple in the passband with an accuracy of 8 bits and reduced the number of taps to 41. Dolecek and Mitra [7] proposed a two-stage CIC-based decimation filter in which a sine-based compensation filter is introduced to decrease the passband drop and a cosine filter is introduced to improve the stopband characteristics. In [8], the

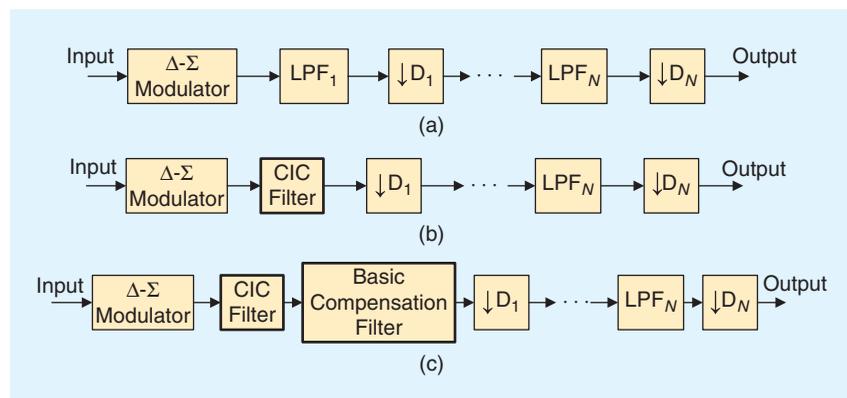


FIGURE 1. A diagram of a CIC filter and a basic compensation filter in Δ - Σ ADC. (a) A basic Δ - Σ ADC. (b) A Δ - Σ ADC in which CIC filter is introduced in the first stage. (c) A Δ - Σ ADC in which both CIC filter and compensation filter are introduced in the first stage.

Digital Object Identifier 10.1109/MSP.2016.2550011
Date of publication: 1 July 2016

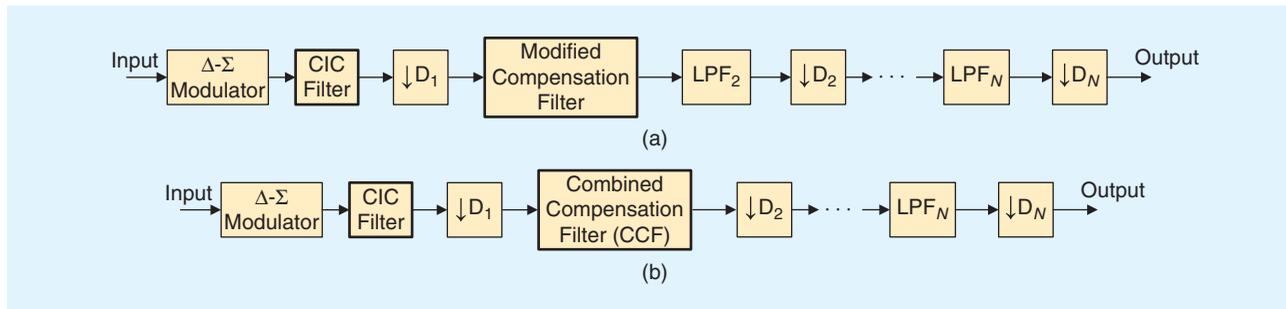


FIGURE 2. A diagram of modified and combined compensation filters in Δ - Σ ADC. (a) A Δ - Σ ADC with compensation filter moved after the first stage. (b) A Δ - Σ ADC with LPF_2 in the second stage merged into the modified compensation filter.

generalized comb filter design methods based on maximally flat, least-square, and minimum-maximum criteria were discussed. The methods were simple and can be applied to narrow- and wide-passband compensation.

Although the aforementioned methods may sometimes achieve a good compensation effect, there is another problem that has not been taken into account, i.e., the spectrum aliasing after downsampling. Intuitively, in a Δ - Σ ADC, a compensation filter is usually introduced after the CIC filter and followed by decimators, as shown in Figure 1(c). It is well known that spectrum aliasing is inevitable when a signal passes through a downsampler [9]. In some applications, aliasing is neglected if the tolerance for the passband ripple is high enough. However, the passband ripple introduced by aliasing is noticeable in high-precision applications. Although the problem can be solved by greatly increasing the attenuation in the stopband, the side effect is obvious, i.e., the filter's order will be increased greatly. In this context, we present here a design method for the FIR compensation filter that trades off computational complexity with compensation performance. This method takes the influence of aliasing into account, and optimizes a cost function that satisfies both constraints in the passband and stopband based on the least squares criterion.

A new design technique for precise FIR compensation filter

In a typical Δ - Σ ADC, the downsampling factor D_1 is relatively large. To reduce the computational complexity, the FIR compensation filter can be

moved to the place after the first decimator, which is renamed as the modified compensation filter as shown in Figure 2(a).

In addition, to further reduce the number of filters, the LPF_2 in the second stage can be merged into the modified compensation filter and be named as the combined compensation filter (CCF), as is shown in Figure 2(b). It is well documented that an effective lowpass filter for a DSM with order L is a cascade of at least three digital CIC filters [4]. For instance, a second-order DSM will incorporate a cascade of at least three digital CIC filters to cancel the quantization noise.

To suppress the inevitable aliasing, some measures should be taken, for example, using the Johnston's technique [9]. Partly motivated by this idea, a new CCF design method is presented next.

Assume that the CIC filter in Figure 2 has a unit impulse response $h_d(n)$ and its transfer function is $H_d(e^{j\omega})$; the FIR CCF has real unit impulse response $h(n)$ and its transfer function is $H(e^{j\omega})$. To obtain a perfect compensation, $H(e^{j\omega})$ should satisfy the following condition:

$$\begin{cases} |H_d(e^{j\omega})||H(e^{j\omega})|=1, & 0 < \omega < \omega_p, \\ |H(e^{j\omega})|=0, & \omega_s < \omega < \pi, \end{cases} \quad (1)$$

where ω is the angular frequency and ω_p and ω_s are the passband and stopband cutoff frequencies, respectively, usually $\omega_p < \omega_s$.

The frequency response $H(e^{j\omega})$ in (1) is not achievable in practice because it constitutes an ideal filter. As an alternative scheme, the approximate $H(e^{j\omega})$, can be obtained by minimizing the following cost function J :

$$J = \alpha \int_0^{\omega_p} [|H(e^{j\omega})H_d(e^{j\omega})|^2 - 1]^2 d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} |H(e^{j\omega})|^2 d\omega, \quad (2)$$

where α is a weight coefficient to weigh the frequency response magnitude error between passband and stopband. Usually $0 < \alpha < 1$.

Assume that $h(n)$ has even symmetry by design and its order is $2M + 1$, where M is a positive integer. We have

$$\begin{aligned} H(e^{j\omega}) &= \sum_{n=-M}^M h(n)e^{-j\omega n} \\ &= h(0) + 2 \sum_{n=1}^M h(n)\cos(n\omega). \end{aligned} \quad (3)$$

Assume that $x'(n)$ is the D -times decimation of a signal $x(n)$. According to the theory of signal decimation [10], its discrete-time Fourier transform (DTFT) is

$$X'(e^{j\omega}) = \frac{1}{D} \sum_{k=0}^{D-1} X(e^{j(\omega - 2k\pi)/D}), \quad (4)$$

where $X(e^{j\omega})$ is the DTFT of $x(n)$. Then, when downsampling and aliasing are considered, the cost function in (2) shall be rewritten as shown in (5):

$$J = a \int_0^{\omega_p} \left[\left| \frac{1}{D} \sum_{k=0}^{D-1} H(e^{j(\omega - 2k\pi)/D}) H_d(e^{j(\omega - 2k\pi)/D}) \right|^2 - 1 \right]^2 d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} |H(e^{j\omega})|^2 d\omega. \quad (5)$$

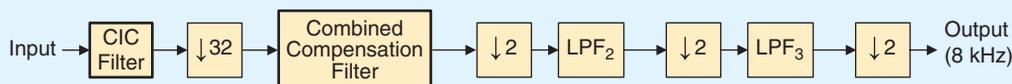


FIGURE 3. A filter design example.

There are several methods available for solving an optimization problem like (5), including the steepest descent, the Newton, and the conjugate gradient methods [12]. Because filter design is an offline computation procedure and convergence time is not a bottleneck problem, we can employ the steepest descent method with a fixed step size μ to minimize (5).

Design example

A series of filters for downsampling Δ - Σ ADCs are shown in Figure 3. Assume that 1) the Δ - Σ modulator is second order and 2) the CIC filter works at the sampling rate of 2,048 kHz and the sampling rate of the final digital output is 8 kHz. It means that the final output signal is in the frequency band 0–4 kHz. Theoretically, when downsampling from 2,048 kHz to 8 kHz, there are many options for D_1, D_2, \dots, D_N . Take three stages ($N = 3$) for example: $D_1 = 8, D_2 = 8$ and $D_3 = 4$; $D_1 = 16, D_2 = 2$ and $D_3 = 8$; and $D_1 = 2, D_2 = 4$ and $D_3 = 32$, and so on. Usually, D_1 is much greater than the rest to reduce the computational complexity, and $D_2 = D_3 = \dots = D_N = 2$ are preferred for the following stages to make the structure regular, simplify the IC chip design, reduce the computational complexity, and save memory space.

The typical transfer function of a CIC is [4],

$$H_{CIC}(z) = \left(\frac{1}{N} \cdot \frac{1 - z^{-N}}{1 - z^{-1}} \right)^L, \quad (6)$$

where N is the order of a single comb filter section and L is the number of sec-

tions. For the design example, N is 32 matching to $D_1 = 32$, and L should be at least three as the aforementioned conclusion in [4]. Thus, the CIC filter transfer function, i.e., $H_d(z)$, is chosen as

$$H_d(z) = H_{CIC}(z) = \left(\frac{1}{32} \cdot \frac{1 - z^{-32}}{1 - z^{-1}} \right)^3, \quad (7)$$

and the frequency response magnitude of the CIC filter is

$$|H_d(e^{j\omega})| = \left| \left(\frac{\sin \frac{32\omega}{2}}{32 \sin \frac{\omega}{2}} \right)^3 \right|. \quad (8)$$

Considering that the compensation filter is moved after the decimation with a factor of 32, the compensation target frequency response magnitude can be expressed as

$$|H_d(e^{j\omega})| = \left| \left(\frac{\sin \frac{\omega}{2}}{32 \sin \frac{\omega}{64}} \right)^3 \right|. \quad (9)$$

For the combined compensation filter $h(n)$, M is set as six, which means that the length of $h(n)$ is 13. Let $h_0(n)$ be the initial solution to this design example, i.e., the initial values of low-pass filter coefficients $h(n)$, obtained from MATLAB with the “fdatool” function. In fact, the initial solution $h_0(n)$ may be simply set as random numbers or zeros. However, a set of good initial coefficient values can really increase the speed to reach the final solution. An example of $h_0(n)$ is $h_0(n) = \{2.7031318e-13,$

$0.0137136, -8.2885986e-13, -0.0628073, 1.4494811e-12, 0.3006935, 0.5, 0.3006935, 1.4494811e-12, -0.0628073, -8.2885986e-13, 0.0137136, 2.7031318e-13\}$.

Without compensation, the maximum ripple value of the CIC filter in the output signal frequency band (0–4 kHz) is about 0.167 dB, while the ideal frequency response magnitude in this band should be 0 dB. To achieve the desired CCF, four cost functions are tested respectively in the following sections. During optimization, the steepest descent iterative process is stopped when the absolute value in the difference for values of J between the m th and the $(m + 1)$ th iterations is lower than a small and positive number ϵ , e.g., $\epsilon = 10^{-14}$.

Basic cost function without considering decimation

According to (9), $|H_d(e^{j\omega})|$ is positive and real, and in the passband ($0 < \omega < \omega_p$), $|H_d(e^{j\omega})| \approx 1$. Thus, when $|H_d(e^{j\omega})| |H(e^{j\omega})| = 1$ is considered, setting $H(e^{j\omega}) = 1$ is a reasonable configuration in the passband ($0 < \omega < \omega_p$). Usually, before and after optimizing, only tiny variations occur in $h(n)$, and then we still have $H(e^{j\omega}) \approx 1$. This means that $H(e^{j\omega})$ and $|H_d(e^{j\omega})|$ are both positive and real values in the passband. When considering that $H(e^{j\omega})$ should be $1/|H_d(e^{j\omega})|$ in the passband and $H(e^{j\omega})$ is also real value in the stopband ($\omega_s < \omega < \pi$), (2) can be simplified directly as

$$J = \alpha \int_0^{\omega_p} \left[H(e^{j\omega}) - \frac{1}{|H_d(e^{j\omega})|} \right]^2 d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} H^2(e^{j\omega}) d\omega. \quad (10)$$

$$\frac{\partial J}{\partial h(n)} = \begin{cases} \alpha \int_0^{\omega_p} \left\{ 2 \left[H(e^{j\omega}) - \frac{1}{|H_d(e^{j\omega})|} \right] \right\} d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} 2H(e^{j\omega}) d\omega, & n = 0, \\ \alpha \int_0^{\omega_p} \left\{ 2 \left[H(e^{j\omega}) - \frac{1}{|H_d(e^{j\omega})|} \right] \cdot 2 \cos(n\omega) \right\} d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} [2H(e^{j\omega}) \cdot 2 \cos(n\omega)] d\omega, & n \neq 0. \end{cases} \quad (11)$$

Table 1. The ripple with a CCF designed from (11).

α	0.1	0.5	0.9	0.95	0.99	0.993	0.994	0.995
Maximum ripple before decimation ($\times 10^{-3}$ dB)	3.0580	1.2162	0.9434	9.3753	9.4632	9.4692	9.4693	9.4669
Maximum ripple after decimation ($\times 10^{-3}$ dB)	2.9133	0.9995	0.8291	7.0402	5.2728	4.6958	4.5591	4.8594

Table 2. The ripple with a CCF designed from (13).

α	0.1	0.5	0.9	0.95	0.98	0.984	0.985	0.986	0.99
Maximum ripple before decimation ($\times 10^{-4}$ dB)	17.7504	9.5646	8.5419	8.5612	8.5695	8.5666	8.5651	8.5632	8.5477
Maximum ripple after decimation ($\times 10^{-4}$ dB)	15.7563	8.7989	5.9189	5.3798	4.5735	4.2597	4.2245	4.3038	4.8196

The gradient of J used in the steepest descent method is shown in (11) at the bottom of the previous page.

The maximum ripple values of pass-band before and after decimation are shown in Table 1 for different values of α during the CCF design. From Table 1, the best compensation effect after decimation is achieved when $\alpha = 0.994$, $\mu = 0.035$, after about 40, 000 iterations (see the section “More discussion about parameters α , μ , and b ”) and the obtained CCF coefficients $h(n) = \{0.0019784, 0.0094407, -0.00939767337644, -0.066953, 0.0041648, 0.3075236, 0.5065062, 0.3075236, 0.0041648, -0.066953, -0.00939767337644, 0.0094407, \text{ and } 0.0019784\}$.

Direct cost function without considering decimation

Considering that $H(e^{j\omega})$ is real valued, the direct version of (2) can be written as

$$J = \alpha \int_0^{\omega_p} \{ [H(e^{j\omega}) | H_d(e^{j\omega})|^2 - 1]^2 d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} H^2(e^{j\omega}) d\omega, \quad (12)$$

and the gradient of J from (12) is shown in (13) at the bottom of the page.

Table 2 shows the compensation results based on (12). Obviously, the best compensation effect after decimation is obtained when $\alpha = 0.985$. In this case, the algorithm converges after about 40, 000 iterations with step size $\mu = 0.012$, the CCF coefficients are $h(n) = \{0.0020063, 0.0094647, -0.009458, -0.0670014, 0.0041785, 0.3075506, 0.5065425, 0.3075506, 0.0041785, -0.0670014, -0.009458, 0.0094647, 0.0020063\}$.

Direct cost function considering decimation

In most cases, to reduce the computational complexity, the downsampling factor D

following the compensation filter is two. When the decimator and aliasing are considered, based on $D = 2$ and referring to Johnston’s cost function [9], the cost function in (5) can be rewritten as (14), and then the gradient of J from (14) is shown in (15), both shown in the box at the bottom of the page.

Table 3 shows the maximum ripples for the compensated filters. When $\alpha = 0.986$, $\mu = 0.015$, the compensation effect is the best after about 30,000 iterations, with the CCF being $h(n) = \{0.0020086, 0.0094673, -0.0094629, -0.067007, 0.0041794, 0.3075538, 0.5065457, 0.3075538, 0.0041794, -0.067007, -0.0094629, 0.0094673, 0.0020086\}$.

Weighted cost function considering decimation

Usually, the total FRM is not equiripple in the passband, especially, the ripple around 4 kHz is larger than the one around the 0–3 kHz band. Because of this, we add a

$$\frac{\partial J}{\partial h(n)} = \begin{cases} \alpha \int_0^{\omega_p} \{ 2[(H(e^{j\omega}) | H_d(e^{j\omega})|^2 - 1) \cdot [2H(e^{j\omega}) | H_d(e^{j\omega})|^2]] d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} 2H(e^{j\omega}) d\omega, & n = 0, \\ \alpha \int_0^{\omega_p} \{ 2[(H(e^{j\omega}) | H_d(e^{j\omega})|^2 - 1) \cdot [2H(e^{j\omega}) | H_d(e^{j\omega})|^2] \cdot 2 \cos(n\omega)] d\omega \\ + (1 - \alpha) \int_{\omega_s}^{\pi} [2H(e^{j\omega}) \cdot 2 \cos(n\omega)] d\omega, & n \neq 0. \end{cases} \quad (13)$$

$$J = \alpha \int_0^{\omega_p} [(H(e^{j\omega}) | H_d(e^{j\omega})|^2 + (H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)})|^2 - 1)^2 d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} H^2(e^{j\omega}) d\omega, \quad (14)$$

$$\frac{\partial J}{\partial h(n)} = \begin{cases} \alpha \int_0^{\omega_p} \{ 2[(H(e^{j\omega}) | H_d(e^{j\omega})|^2 + (H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)})|^2 - 1) \cdot [2H(e^{j\omega}) | H_d(e^{j\omega})|^2 + 2H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)})|^2]] d\omega \\ + (1 - \alpha) \int_{\omega_s}^{\pi} 2H(e^{j\omega}) d\omega, & n = 0, \\ \alpha \int_0^{\omega_p} \{ 2[(H(e^{j\omega}) | H_d(e^{j\omega})|^2 + (H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)})|^2 - 1) \cdot [2H(e^{j\omega}) | H_d(e^{j\omega})|^2 \cdot 2 \cos(n\omega) + 2H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)})|^2 \cdot 2 \cos(n(\omega - \pi))]] d\omega \\ + (1 - \alpha) \int_{\omega_s}^{\pi} [2H(e^{j\omega}) \cdot 2 \cos(n\omega)] d\omega, & n \neq 0. \end{cases} \quad (15)$$

Table 3. The ripple with a CCF designed from (15).

α	0.1	0.5	0.9	0.95	0.98	0.985	0.986	0.987	0.99
Maximum ripple before decimation ($\times 10^{-4}$ dB)	16.7026	9.4152	8.4530	8.4768	8.4881	8.4847	8.4832	8.4812	8.4697
Maximum ripple after decimation ($\times 10^{-4}$ dB)	14.8461	8.6151	5.8722	5.3529	4.6136	4.2359	4.1974	4.2809	4.6547

Table 4. The ripple with a CCF designed from (17).

α	0.1	0.5	0.9	0.91	0.917	0.918	0.919	0.92	0.95
Maximum ripple before decimation ($\times 10^{-4}$ dB)	24.2787	7.7253	7.1145	7.1078	7.1031	7.1024	7.1017	7.1010	7.0778
Maximum ripple after decimation ($\times 10^{-4}$ dB)	24.0337	6.5592	3.8409	3.8225	3.8085	3.8069	3.8146	3.8226	4.2369

weighting function that depends on frequency and is independent of α , so as to equalize the ripple in the passband, in other words, to make the maximum passband ripple as small as possible after compensation and decimation. Therefore, the cost function in (14) is modified as (16), where $w(\omega)$ is the weighting function, and the gradient of J from (16) is shown in (17), both of which are shown in the box at the bottom of the page.

In this example, the weighting function is chosen as $w(\omega) = b^\omega$, where b

is set as 47 (further discussion about b appears in the section “More discussion about parameters α , μ , and b ”). Table 4 presents the maximum ripples when the CCFs are designed with the cost function (16). When $\alpha = 0.918$ and $\mu = 0.0009$, the compensation effect after decimation is the best after about 186,000 iterations, with the CCF being $h(n) = \{0.00207, 0.0094925, -0.0095819, -0.0670571, 0.0041985, 0.3075818, 0.5066276, 0.3075818, 0.0041985, -0.0670571, -0.0095819, 0.0094925, 0.00207\}$.

According to the data presented in Table 4, it is clear that the cost function in the section “Weighted Cost Function Considering Decimation” gives a better result than those in the sections “Basic Cost Function Without Considering Decimation,” “Direct Cost Function Without Considering Decimation,” and “Direct Cost Function Considering Decimation.” The frequency response magnitude of a FIR filter designed with this weighted cost function is shown in Figure 4, and the compensation effect is shown in Figure 5.

From Figure 5 we can see that the frequency response magnitude curves of the original CIC filter and the designed CCF frequency response magnitude curves are almost symmetrical about the horizontal axis (magnitude = 0 dB), and the total frequency response magnitude is flat in the band from 0 to 4 kHz. As the frequency response magnitudes of the compensated CIC filter before and after decimation are almost the same, Figure 6 presents the details of the compensation effect before and after decimation in the band from 0 to 4 kHz. We can easily find two facts: 1) the peak-to-peak ripple of the compensated

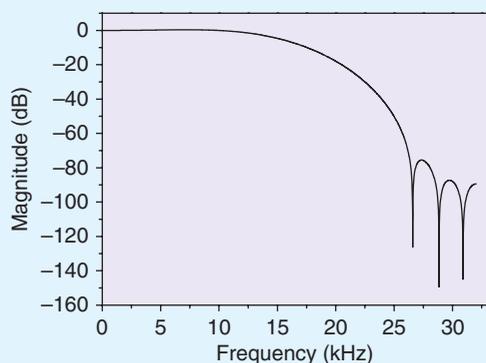


FIGURE 4. The FRM of CCF in the section “Weighted Cost Function Considering Decimation.”

$$J = \alpha \int_0^{\omega_p} \{ [(H(e^{j\omega}) | H_d(e^{j\omega})) |^2 + (H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)}) |^2 - 1)]^2 w(\omega) \} d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} H^2(e^{j\omega}) d\omega, \quad (16)$$

$$\frac{\partial J}{\partial h(n)} = \begin{cases} \alpha \int_0^{\omega_p} \{ 2w(\omega) [(H(e^{j\omega}) | H_d(e^{j\omega})) |^2 + (H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)}) |^2 - 1) \cdot [2H(e^{j\omega}) | H_d(e^{j\omega})) |^2 + 2H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)}) |^2] \} d\omega + (1 - \alpha) \int_{\omega_s}^{\pi} 2H(e^{j\omega}) d\omega, & n = 0, \\ \alpha \int_0^{\omega_p} \{ 2w(\omega) [(H(e^{j\omega}) | H_d(e^{j\omega})) |^2 + (H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)}) |^2 - 1) \cdot [2H(e^{j\omega}) | H_d(e^{j\omega})) |^2 \cdot 2 \cos(n\omega) + 2H(e^{j(\omega-\pi)}) | H_d(e^{j(\omega-\pi)}) |^2 \cdot 2 \cos(n(\omega - \pi))] \} d\omega & n \neq 0. \\ + (1 - \alpha) \int_{\omega_s}^{\pi} [2H(e^{j\omega}) \cdot 2 \cos(n\omega)] d\omega, & \end{cases} \quad (17)$$

CIC filter after decimation is smaller than that before decimation and 2) compared with the best results in the section “Direct Cost Function Considering Decimation,” after decimation, the maximum ripple of the compensated CIC filter in the passband is reduced to about 3.8×10^{-4} dB.

More discussion about parameters α , μ , and b

The purpose of the weight coefficient α in (11), (13), (15), and (17) is to trade off the performance between compensation in passband and attenuation in stopband. Usually, the initial solution of a lowpass filter $h(n)$ designed by the MATLAB ‘`fdatool`’ function (or other computer-aided design software) has a good attenuation performance in stopband, therefore, the CCF design should place emphasis on the compensation performance, in other words, α should be closer to 1, for possible initial choices in the range $0 < \alpha < 1$. To further determine the value, a standard binary search can be employed. According to our filter design experience, several attempts (usually, fewer than ten) are enough to find out an acceptable α , as seen in Tables 1–4.

Generally, the step size μ used in the steepest descent algorithm affects the convergence time for the iterative process. For a given cost function, a proper μ should be considered. A standard binary search can still be used to find out a proper step size μ . In the initial attempt, μ should be small enough to guarantee the convergence of the iterative process. When considering that the filter design is an offline task, the result precision is more important than the time to find the solution. All of these reasons point toward choosing a small μ just like in the section “Weighted Cost Function Considering Decimation.”

When choosing a weighting function in the section “Weighted Cost Function Considering Decimation,” the exponential function $w(\omega) = b^\omega$ was chosen because of its simplicity. How should one determine b ? For example, in the best results of Table 3, there are two local maxima: 2.8325×10^{-4} dB at 0 Hz and 4.1974

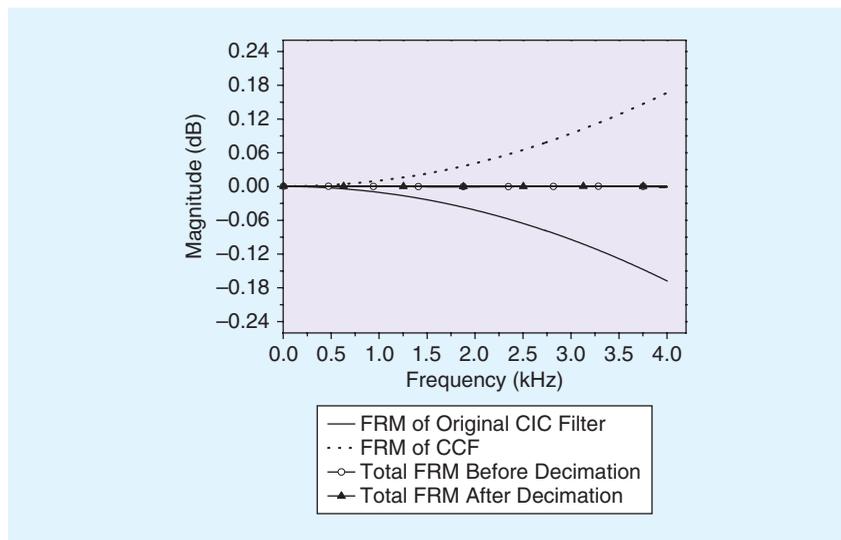


FIGURE 5. The FRMs of CIC filter, CCF, and total FRM of compensated CIC filter before and after decimation in section “Weighted Cost Function Considering Decimation.”

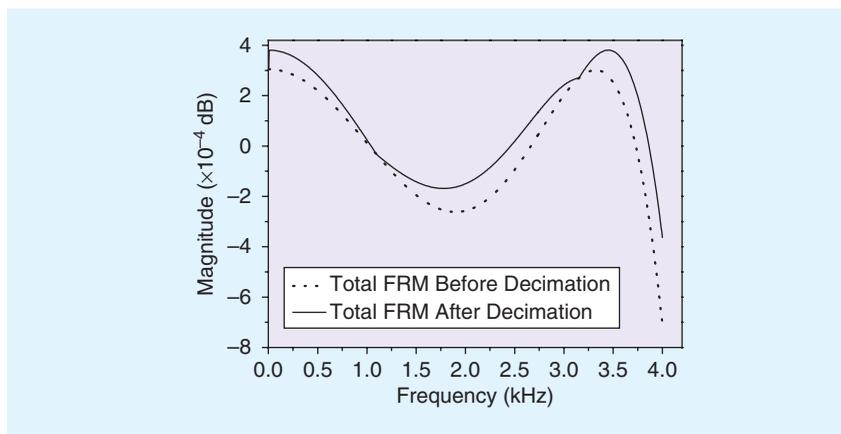


FIGURE 6. Total FRMs of compensated CIC filter before and after decimation in the section “Weighted Cost Function Considering Decimation.”

$\times 10^{-4}$ dB at 3.43 KHz, i.e., 2.8325×10^{-4} dB at $\omega = 0$ rad and 4.1974×10^{-4} dB at $\omega = 0.3367$ rad. To make the two local maxima as same as possible, we set $w(0.3367)/w(0) = b^{0.3367} = (4.1974 \times 10^{-4}/2.8325 \times 10^{-4})$ $\beta \approx 1.4819$ β , where $\beta = 2.5$ is an empirical constant. Then we obtain $b \approx 48.9$. Finally, after some fine-tuning, b is set as 47, and now the two local maxima are almost same. Without carrying out the design method in the section “Direct Cost Function Considering Decimation,” the local maxima are not available. In this case, since b is always greater than 1, we have to search a proper b with the standard binary search. Incidentally, other weighting functions with a

similar weighting behavior can also be applied here.

Note that, α , μ , and b can be determined independently because there is almost no coupling among them. If a priority order is asked for searching or determining their values, our advice is first μ , then α , and last b .

Conclusions

In this article, a novel combined compensation filter design method is presented for the compensation of the CIC filter in a Δ - Σ ADC. It provides the possibility to compensate the CIC filter with a relatively low-order FIR filter without the loss of performance. Compared with those results without the compensation filter, a

low-order, well-designed compensation filter designed by the proposed method can decrease greatly the maximum ripple in the passband. Considering the decimator and when a weighting cost function is used, the maximum frequency response magnitude passband ripple of the cascaded integrator-comb filters and the FIR combined compensation filter is less than 0.0004 dB even if the combined compensation filter order is only 13.

Acknowledgments

We would like to thank the reviewers for their valuable comments, which were very helpful in improving this article. This work was partly supported by the National Natural Science Foundation of China (61172107, 61172110), the National High Technology Research and Development Program (863 Program) of China (2015AA016306), the Major Projects in Liaoning Province Science and Technology Innovation (201302001), and the Fundamental Research Funds for the Central Universities of China (DUT13LAB06).

Authors

Zhe Chen (zhechen@dlut.edu.cn) is an associate professor in the School of Information and Communication Engineering, Dalian University of Technology, China.

Shuwen Wang (wsw@mail.dlut.edu.cn) is a graduate student in the School of Information and Communication Engineering, Dalian University of Technology, China.

Fuliang Yin (flyin@dlut.edu.cn) is a professor in the School of Information and Communication Engineering, Dalian University of Technology, China, and the director of Liaoning province key laboratory for digital media processing and transmission, China.

References

- [1] D. A. Johns and K. Martin, *Analog Integrated Circuit Design*. New York: Wiley, 1997.
- [2] R. E. Crochiere and L. R. Rabiner, "Optimum FIR digital filter implementations for decimation, interpolation, and narrowband filtering," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 23, no. 5, pp. 444–456, 1975.

[3] R. E. Crochiere and L. R. Rabiner, "Further considerations in the design of decimators and interpolators," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 24, no. 4, pp. 296–311, 1976.

[4] S. Ren, R. Siferd, R. Blumgold, and R. Ewing, "Hardware efficient FIR compensation filter for delta sigma modulator analog to digital converters," in *Proc. IEEE 48th Midwest Symp. Circuits and Systems*, 2005, pp. 1514–1517.

[5] A. Y. Kwentus, Z. Jiang, and A. N. Willson, "Application of filter sharpening to cascaded integrator-comb decimation filters," *IEEE Trans. Signal Processing*, vol. 45, no. 2, pp. 457–467, 1997.

[6] H. J. Oh, K. Sunbin, G. Choi, and Y. H. Lee, "On the use of interpolated second-order polynomials for efficient filter design in programmable downconversion," *IEEE J. Select. Areas Commun.*, vol. 17, no. 4, pp. 551–560, 1999.

[7] G. J. Dolecek and S.K. Mitra, "A new two-stage cic-based decimation filter," in *Proc. 5th Int. Symp. Image and Signal Processing and Analysis*, Istanbul, Turkey, 2007, pp. 218–223.

[8] A. Fernandez-Vazquez and G. Jovanovic-Dolecek, "A general method to design GCF compensation filter," *IEEE Trans. Circuits Systems II*, vol. 5, no. 5, pp. 409–413, 2009.

[9] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1986.

[10] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 2007.

[11] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice Hall, 2001.

[12] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. New York: Wiley, 2000.



Article Series on Signal Processing in Hands-on and Design Projects Call for Information on Student Projects

An article series being developed by a guest editor team for IEEE Signal Processing Magazine (SPM) seeks to provide a platform to share experiences and best practices on hands-on and design trainings that incorporate signal and information processing. SPM invites inputs on recent student projects of this type.

What? Input can be made by faculty or industry professionals who have served as mentors for the student projects; input from students is welcome, but endorsement by a faculty/mentor should be included with the submission. Due to space limit, project information will not be published in full in the SPM; instead, the guest editors will select representative projects and compile an article to include summaries or highlights of the projects, and published in SPM with an acknowledgement of project authors and mentors.

How? Interested project authors and mentors should post their project information as a PDF file on the online repository of the IEEE Signal Processing Society, SigPort.org, under the "Event/Theme" category of "SPM Student Design Project Series" using coupon code "spm1227" and their IEEE web account. The submission should have a cover page that includes the project name, institution, name and email contact of project author(s) and mentor(s), project date/duration, and URL if available. A summary up to 3 pages should describe succinctly the project motivation and main elements as well as outcome and notably highlights. At least one image providing a visual summary or highlight of the project should be included; more relevant images/diagram/photos are welcome; an informative thumbnail image should be included on the SigPort platform. A full project report can be appended as one PDF file, or a URL of the report given.

Guest Editor Team: Hana Godrich, Arye Nehorai, Ali Tajer, Maria Sabrina Greco, and Changshui Zhang

When? Submissions are evaluated on a rolling basis by **31 July 2016**. Early submission is encouraged.

Digital Object Identifier 10.1109/MSP.2016.2583386



© GRAPHIC STOCK

ADVERTISING & SALES

The Advertisers Index contained in this issue is compiled as a service to our readers and advertisers: the publisher is not liable for errors or omissions although every effort is made to ensure its accuracy. Be sure to let our advertisers know you found them through IEEE Signal Processing Magazine.

CVR 2

IEEE MDL/Marketing
www.ieee.org/go/trymdl

CVR 4

Mathworks
www.mathworks.com/wireless
+1 508 647 7040

James A. Vick

Sr. Director, Advertising
Phone: +1 212 419 7767;
Fax: +1 212 419 7589
jv.ieeemedia@ieee.org

Marion Delaney

Advertising Sales Director
Phone: +1 415 863 4717;
Fax: +1 415 863 4717
md.ieeemedia@ieee.org

Mark David

Sr. Manager Advertising & Business Development
Phone: +1 732 465 6473
Fax: +1 732 981 1855
m.david@ieee.org

Mindy Belfer

Advertising Sales Coordinator
Phone: +1 732 562 3937
Fax: +1 732 981 1855
m.belfer@ieee.org

PRODUCT ADVERTISING

MIDATLANTIC

Lisa Rinaldo
Phone: +1 732 772 0160;
Fax: +1 732 772 0164
lr.ieeemedia@ieee.org
NY, NJ, PA, DE, MD, DC, KY, WV

NEW ENGLAND/ SOUTH CENTRAL/ EASTERN CANADA

Jody Estabrook
Phone: +1 774 283 4528;
Fax: +1 774 283 4527
je.ieeemedia@ieee.org
ME, VT, NH, MA, RI, CT, AR, LA, OK, TX
Canada: Quebec, Nova Scotia, Newfoundland, Prince Edward Island, New Brunswick

SOUTHEAST

Cathy Flynn
Phone: +1 770 645 2944;
Fax: +1 770 993 4423
cf.ieeemedia@ieee.org
VA, NC, SC, GA, FL, AL, MS, TN

MIDWEST/ CENTRAL CANADA

Dave Jones
Phone: +1 708 442 5633;
Fax: +1 708 442 7620
dj.ieeemedia@ieee.org
IL, IA, KS, MN, MO, NE, ND, SD, WI, OH
Canada: Manitoba, Saskatchewan, Alberta

MIDWEST/ ONTARIO, CANADA

Will Hamilton
Phone: +1 269 381 2156;
Fax: +1 269 381 2556
wh.ieeemedia@ieee.org
IN, MI, Canada: Ontario

WEST COAST/ MOUNTAIN STATES/ WESTERN CANADA

Marshall Rubin
Phone: +1 818 888 2407;
Fax: +1 818 888 4907
mr.ieeemedia@ieee.org
AZ, CO, HI, NM, NV, UT, AK, ID, MT, WY, OR, WA, CA. Canada: British Columbia

EUROPE/AFRICA/ MIDDLE EAST

**ASIA/FAR EAST/
PACIFIC RIM**
Michael O'Kane
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
mo.ieeemedia@ieee.org
Europe, Africa, Middle East
Asia, Far East, Pacific Rim, Australia, New Zealand

RECRUITMENT ADVERTISING

MIDATLANTIC

Lisa Rinaldo
Phone: +1 732 772 0160;
Fax: +1 732 772 0164
lr.ieeemedia@ieee.org
NY, NJ, CT, PA, DE, MD, DC, KY, WV

NEW ENGLAND/ EASTERN CANADA

Liza Reich
Phone: +1 212 419 7578;
Fax: +1 212 419 7589
e.reich@ieee.org
ME, VT, NH, MA, RI, Canada: Quebec, Nova Scotia, Prince Edward Island, Newfoundland, New Brunswick

SOUTHEAST

Cathy Flynn
Phone: +1 770 645 2944;
Fax: +1 770 993 4423
cf.ieeemedia@ieee.org
VA, NC, SC, GA, FL, AL, MS, TN

MIDWEST/ SOUTH CENTRAL/ CENTRAL CANADA

Darcy Giovino
Phone: +224 616 3034;
Fax: +1 847 729 4269
dg.ieeemedia@ieee.org
AR, IL, IN, IA, KS, LA, MI, MN, MO, NE, ND, SD, OH, OK, TX, WI. Canada: Ontario, Manitoba, Saskatchewan, Alberta

WEST COAST/ SOUTHWEST/ MOUNTAIN STATES/ASIA

Tim Matteson
Phone: +1 310 836 4064;
Fax: +1 310 836 4067
tm.ieeemedia@ieee.org
AZ, CO, HI, NV, NM, UT, CA, AK, ID, MT, WY, OR, WA. Canada: British Columbia

EUROPE/AFRICA/ MIDDLE EAST

Michael O'Kane
Phone: +44 1875 825 700;
Fax: +44 1875 825 701
mo.ieeemedia@ieee.org
Europe, Africa, Middle East



DATES AHEAD

Please send calendar submissions to:

Jessica Barragué

E-mail: j.barrague@ieee.org

2016

JULY

IEEE Ninth Sensor Array and Multichannel Signal Processing Workshop (SAM)

10–13 July, Rio de Janeiro, Brazil.

General Chairs: Rodrigo C. de Lamare and Martin Haardt

URL: <http://delamare.cetuc.puc-rio.br/sam2016/index.html>**IEEE 12th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)**

11–12 July, Bordeaux, France.

General Chair: Yannick Berthoumieu

URL: <http://ivmsp2016.org/>**IEEE International Conference on Multimedia and Expo (ICME)**

11–15 July, Seattle, Washington, USA.

General Chairs: Tsuhan Chen, Ming-Ting Sun, and Cha Zhang

URL: <http://www.icme2016.org/>

AUGUST

IEEE 13th International Conference on Advanced Video and Signal Based Surveillance (AVSS)

23–26 August, Colorado Springs, Colorado, USA.

General Cochairs: Terry Boult and Ram Nevatia

URL: <http://avss2016.org/>**IEEE 24th European Signal Processing Conference (EUSIPCO)**

29 August–2 September, Budapest, Hungary.

General Chair: Lajos Hanzo

URL: <http://www.eusipco2016.org/>

SEPTEMBER

IEEE International Workshop on Machine Learning for Signal Processing (MLSP)

13–16 September, Salerno, Italy.

Fourth International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar, and Remote Sensing (CoSeRa)

19–22 September, Aachen, Germany.

Digital Object Identifier 10.1109/MSP.2016.2547162
Date of publication: 1 July 2016

©ISTOCKPHOTO.COM/TONDA

ICIP 2016 will be held in Phoenix, Arizona, United States, 25–28 September, at the Phoenix Convention Center.

IEEE International 18th International Workshop on Multimedia Signal Processing (MMSP)

20–23 September, Montréal, Québec, Canada.

Sensor Signal Processing for Defence (SSPD)

22–23 September, Edinburgh, Great Britain.

URL: <http://www.sspd.eng.ed.ac.uk>**IEEE International Conference on Image Processing (ICIP)**

25–28 September, Phoenix, Arizona, USA.

General Chair: Lina Karam

URL: <http://www.ieeeicip2016.org>

NOVEMBER

50th Annual Asilomar Conference on Signals, Systems, and Computers (ASILOMAR)

6–9 November, Pacific Grove, California, USA.

General Chair: Philip Schniter

URL: <http://www.asilomarsseconf.org/>

DECEMBER

Picture Coding Symposium (PCS)

4–7 December, Nuremberg, Germany.

General Chair: André Kaup

URL: <http://www.pcs2016.com/>**Eighth IEEE International Workshop on Information Forensics and Security (WIFS)**

5–7 December, Abu Dhabi, UAE.

General Chairs: Ernesto Damiani and Nasir Memon

URL: <http://wifs2016.mdabaie.com/>**IEEE Global Conference on Signal and Information Processing (GlobalSIP)**

7–9 December, Greater Washington, D.C., USA.

General Chairs: Zhi Tian and Brian Sadler

URL: <http://2016.ieeeglobalsip.org>**IEEE World Forum on Internet of Things (WF-IoT)**

12–14 December, Reston, Virginia, USA.

General Chairs: Geoff Mulligan and Latif Ladid

URL: <http://wfiot2016.ieee-wf-iot.org/>**IEEE Spoken Language Technology Workshop (SLT)**

13–16 December, San Juan, Puerto Rico.

Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)

13–16 December, Jeju, South Korea.

2017

MARCH

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)

5–9 March, New Orleans, Louisiana, USA.

General Chair: Magdy Bayoumi

URL: <http://www.ieee-icassp2017.org/>

APRIL

IEEE International Symposium on Biomedical Imaging (ISBI)

18–21 April, Melbourne, Australia

URL: <http://biomedicalimaging.org/2017/>

SEPTEMBER

IEEE International Conference on Image Processing (ICIP)

17–20 September, Beijing, China

General Chairs: Xinggang Lin, Anthony Vetro, and Min Wu

URL: <http://2017.ieeeicip.org/>

DECEMBER

17th IEEE International Workshop on Computational Advances in Multisensor Adaptive Processing (CAMSAP)

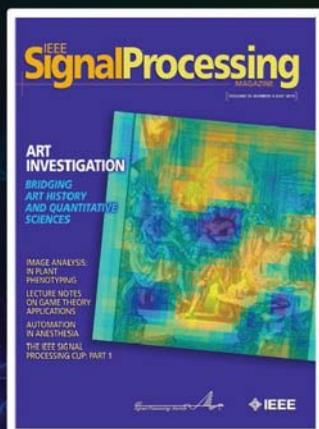
10–13 December, Curacao, Dutch Antilles.

General Chairs: André L.F. de Almeida and Martin Haardt

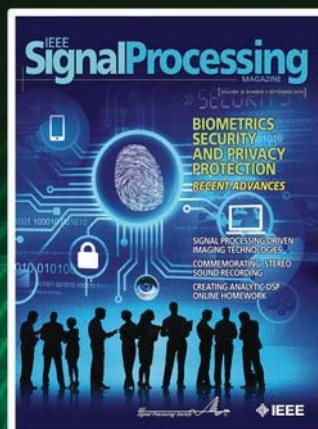
URL: <http://www.cs.huji.ac.il/conferences/CAMSAP17/>

SP

July 2015
Art Investigation



September 2015
Biometrics Security & Privacy



November 2015
Feature Article Collection

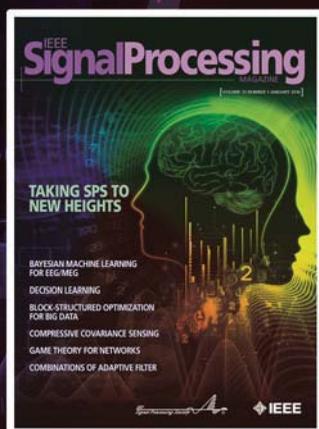


Publish with IEEE Signal Processing Magazine

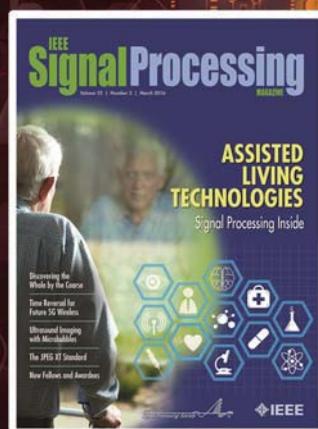
HIGH IMPACT among all electrical engineering publications

REACH a broad signal processing audience worldwide

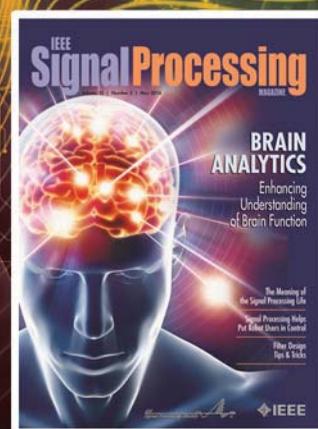
WELCOME proposals for Special Issues and Feature Articles, and contributions to Columns



January 2016
Feature Article Collection



March 2016
Assistive Living Technologies



May 2016
Brain Signal Analytics

MATLAB SPEAKS WIRELESS DESIGN

You can simulate, prototype, and verify wireless systems right in MATLAB. Learn how today's MATLAB supports RF, LTE, WLAN and 5G development and SDR hardware.

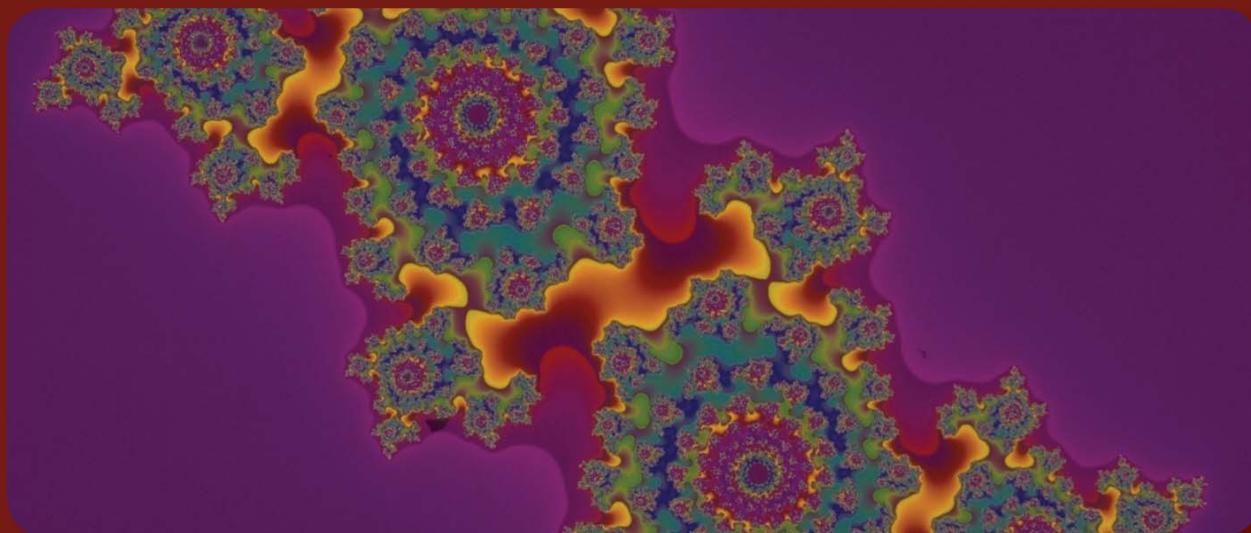
mathworks.com/wireless

IEEE SIGNAL PROCESSING SOCIETY

Content Gazette

JULY 2016

ISSN 2167-5023

**T-SP July 1 2016 Vol. 64 #13**<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7478178>**T-SP July 15 2016 Vol. 64 #14**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7482863>**T-ASLP June 2016 Vol. 24 #6**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7476922>**T-IP June 2016 Vol. 25 #6**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7469921>**T-IFS June 2016 Vol. 11 #6**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7469963>**T-MM June 2016 Vol. 18 #6**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7469930>**T-JSTSP June 2016 Vol. 10 #4**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7469427>**T-SPL June 2016 Vol. 23 #6**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7475984>**T-CI June 2016 Vol. 2 #2**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7464408>**T-SIPN June 2016 Vol. 2 #2**<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7463088>



The **IEEE International Workshop on Information Forensics and Security (WIFS)** is the primary annual event organized by the IEEE Information Forensics and Security (IFS) Technical Committee. Its major objective is to bring together researchers from relevant disciplines to exchange new ideas and the latest results and to discuss emerging challenges in different areas of information security. The 8th edition of WIFS will be held in Abu Dhabi, UAE, from December 4, to December 7, 2016. WIFS 2016 will feature keynote lectures, tutorials, technical & special sessions, and also demo and ongoing works sessions.

Topics of interest include, but are not limited to:

- **Forensics**
Evidence Recovery | Evidence Validation | Media Analysis | Source Attribution | Incident Response | Counter Forensics
- **Biometrics**
- **Steganography and Covert Communications**
- **Multimedia content security**
Cryptography for multimedia | Hashing | Data Hiding
- **Privacy**
Privacy in Data Analytics | Privacy in Internet of Everything
- **Information and System Security**
Authentication | Vulnerability Discovery | Malware Analysis | Anonymity | Cyber-Physical Systems | Mobile and Intelligent Devices | PHY Fingerprinting
- **Hardware security**
Hardware Trojans | PUFs | Anti-Counterfeiting
- **Surveillance and Sousveillance**
Tracking | Person (Re-)Identification | Crowd Analysis | Anti-Surveillance and De-identification | Anti-Sousveillance
- **Network Traffic Analysis**

Submission of papers: Prospective authors are invited to submit full-length, six-page papers, including figures and references. All submitted papers will go through double-blinded peer review process. The WIFS Technical Program Committee will select papers for the formal proceedings based on technical quality, relevance to the workshop, and ability to inspire new research. Accepted papers will be presented in either lecture tracks or poster sessions. Authors of the accepted papers are required to present their papers at the conference. For all questions contact WIFS'16 Technical Program Chairs at tpc@wifs2016.org.

Tutorial Proposals: Up to four tutorials will be scheduled for the first day of the conference, Sunday December 4, 2016. 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 Chair at tutorials@wifs2016.org.

Demo and Ongoing Works Proposal: This session will provide both academic researchers and industrial exhibitors to showcase innovative implementations, systems and technologies demonstrating new ideas in the field. We encourage both the submission of early research prototypes and interesting mature systems. Formal proposals must be accompanied by a description of the work or demo to be presented. All material has to be submitted to the Demo Session Chair at demo@wifs2016.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 2016, subject to space availability and approval by the WIFS Technical Program Chairs. Proposals have to be submitted to the Technical Program Chairs at tpc@wifs2016.org.

Organizing Committee		Important Deadlines	
General Chairs - Ernesto Damiani <i>Khalifa University, UAE</i> - Nasir Memon <i>New York University, USA</i>	Technical Program Chairs - Hüsrev Taha Sencar <i>TOBB University, Turkey</i> - Ahmad-Reza Sadeghi <i>TU Darmstadt, Germany</i>	Special session proposals:	May 16, 2016
		Notification of special session acceptance:	May 27, 2016
Publications Chair - Andrew Ker <i>University of Oxford, UK</i>	Tutorials Chairs - Michail Maniatakos <i>New York University Abu Dhabi, UAE</i>	Tutorial proposals:	May 27, 2016
		Notification of tutorial acceptance:	July 1, 2016
Publicity Chair - Stefan Katzenbeisser <i>TU Darmstadt, Germany</i>	Demo Session Chairs - Yier Jin, <i>University of Central Florida, USA</i> - Khaled Saleh, <i>Khalifa University, UAE</i>	Paper submission deadline:	July 24, 2016
		Demo & on-going work proposals:	August 12, 2016
Finance Chair - Florian Kerschbaum <i>SAP, Germany</i>	Local Arrangements Chair - Özgür Sinanoglu <i>New York University Abu Dhabi, UAE</i>	SPL and TIFS Paper Submissions:	August 26, 2016
		Notification of paper acceptance:	September 25, 2016
		Notification of demo & ongoing works proposals acceptance:	September 25, 2016
		Submission of camera-ready papers:	October 21, 2016
		End of advance registration:	October 21, 2016
		Author's registration deadline:	November 4, 2016

For further details, please visit the WIFS 2016 conference website: <http://www.wifs2016.org>



IEEE TRANSACTIONS ON SIGNAL 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



JULY 1, 2016

VOLUME 64

NUMBER 13

ITPRED

(ISSN 1053-587X)

REGULAR PAPERS

Robust Convex Approximation Methods for TDOA-Based Localization Under NLOS Conditions http://dx.doi.org/10.1109/TSP.2016.2539139	<i>G. Wang, A. M.-C. So, and Y. Li</i>	3281
Iterative Bayesian Reconstruction of Non-IID Block-Sparse Signals http://dx.doi.org/10.1109/TSP.2016.2543208	<i>M. Korki, J. Zhang, C. Zhang, and H. Zayyani</i>	3297
Distributed Inference Over Directed Networks: Performance Limits and Optimal Design http://dx.doi.org/10.1109/TSP.2016.2543209	<i>D. Bajović, J. M. F. Moura, J. Xavier, and B. Sinopoli</i>	3308
Training Design for Channel Estimation in Uplink Cloud Radio Access Networks http://dx.doi.org/10.1109/TSP.2016.2539126	<i>Q. Hu, M. Peng, Z. Mao, X. Xie, and H. V. Poor</i>	3324
Embedded Algorithmic Noise-Tolerance for Signal Processing and Machine Learning Systems via Data Path Decomposition http://dx.doi.org/10.1109/TSP.2016.2546224	<i>S. Zhang and N. R. Shanbhag</i>	3338

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$39.00 (first copy only), non-members \$663.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of \$31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2016 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



Near-Field Acoustic Source Localization and Beamforming in Spherical Harmonics Domain http://dx.doi.org/10.1109/TSP.2016.2543201 ..	
.....	<i>L. Kumar and R. M. Hegde</i> 3351
Deep Sensing for Space-Time Doubly Selective Channels: When a Primary User Is Mobile and the Channel Is Flat Rayleigh Fading http://dx.doi.org/10.1109/TSP.2016.2537276	<i>B. Li, J. Hou, X. Li, Y. Nan, A. Nallanathan, and C. Zhao</i> 3362
Generalized Correntropy for Robust Adaptive Filtering http://dx.doi.org/10.1109/TSP.2016.2539127	<i>B. Chen, L. Xing, H. Zhao, N. Zheng, and J. C. Principe</i> 3376
Optimal Joint User Association and Multi-Pattern Resource Allocation in Heterogeneous Networks http://dx.doi.org/10.1109/TSP.2016.2548998	<i>Q. Kuang, W. Utschick, and A. Dotzler</i> 3388
Multiple-Parameter Discrete Fractional Transform and Its Applications http://dx.doi.org/10.1109/TSP.2016.2544740	<i>X. Kang, R. Tao, and F. Zhang</i> 3402
Type I and Type II Bayesian Methods for Sparse Signal Recovery Using Scale Mixtures http://dx.doi.org/10.1109/TSP.2016.2546231	<i>R. Giri and B. Rao</i> 3418
Faster Learning and Adaptation in Security Games by Exploiting Information Asymmetry http://dx.doi.org/10.1109/TSP.2016.2548987	<i>X. He, H. Dai, and P. Ning</i> 3429
Deep Neural Networks with Random Gaussian Weights: A Universal Classification Strategy? http://dx.doi.org/10.1109/TSP.2016.2546221 ..	<i>R. Giryes, G. Sapiro, and A. M. Bronstein</i> 3444
Power Minimization of a Wireless Sensor Node Under Different Rate Constraints http://dx.doi.org/10.1109/TSP.2016.2548991	<i>J. R. Ayala Solares, L. Sboui, Z. Rezk, and M.-S. Alouini</i> 3458
Distributed Recovery of Jointly Sparse Signals Under Communication Constraints http://dx.doi.org/10.1109/TSP.2016.2548990	<i>S. M. Fosson, J. Matamoros, C. Antón-Haro, and E. Magli</i> 3470
Compressed Sensing with Basis Mismatch: Performance Bounds and Sparse-Based Estimator http://dx.doi.org/10.1109/TSP.2016.2544742 ..	<i>S. Bernhardt, R. Boyer, S. Marcos, and P. Larzabal</i> 3483
Minimax Optimal Sparse Signal Recovery With Poisson Statistics http://dx.doi.org/10.1109/TSP.2016.2529588	<i>M. H. Rohban, V. Saligrama, and D. Motamed Vaziri</i> 3495
Sensor Selection for Estimation with Correlated Measurement Noise http://dx.doi.org/10.1109/TSP.2016.2550005	<i>S. Liu, S. P. Chepuri, M. Fardad, E. Maşazade, G. Leus, and P. K. Varshney</i> 3509
<hr/>	
OVERVIEW ARTICLE	
Recursive Recovery of Sparse Signal Sequences From Compressive Measurements: A Review http://dx.doi.org/10.1109/TSP.2016.2539138	<i>N. Vaswani and J. Zhan</i> 3523

IEEE TRANSACTIONS ON SIGNAL 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



JULY 15, 2016

VOLUME 64

NUMBER 14

ITPRED

(ISSN 1053-587X)

regular papers

Recursive Implementation of the Gaussian Filter Using Truncated Cosine Functions http://dx.doi.org/10.1109/TSP.2016.2549985	D. Charalampidis	3554
Deterministic Compressed Sensing Matrices: Construction via Euler Squares and Applications http://dx.doi.org/10.1109/TSP.2016.2550020	R. R. Naidu, P. Jampana, and C. S. Sastry	3566
Robust Estimation of Structured Covariance Matrix for Heavy-Tailed Elliptical Distributions http://dx.doi.org/10.1109/TSP.2016.2546222 ..	Y. Sun, P. Babu, and D. P. Palomar	3576
A New Theoretical Model for the Pseudo Affine Projection Algorithm for Unity Step Size and Autoregressive Inputs http://dx.doi.org/10.1109/TSP.2016.2549992	M. H. Costa, S. J. M. Almeida, and J. C. M. Bermudez	3591
Information, Estimation, and Lookahead in the Gaussian Channel http://dx.doi.org/10.1109/TSP.2016.2544748	K. Venkat, T. Weissman, Y. Carmon, and S. Shamai	3605

IEEE TRANSACTIONS ON SIGNAL PROCESSING (ISSN 1053-587X) is published semimonthly by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$39.00 (first copy only), nonmembers \$663.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of \$31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2016 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. **Postmaster:** Send address changes to IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



Group Sparsity Based Multi-Target Tracking in Passive Multi-Static Radar Systems Using Doppler-Only Measurements http://dx.doi.org/10.1109/TSP.2016.2552498	<i>S. Subedi, Y. D. Zhang, M. G. Amin, G. Amin, and B. Himed</i>	3619
Communication Using a Large-Scale Array of Ubiquitous Antennas: A Geometry Approach http://dx.doi.org/10.1109/TSP.2016.2544738 ..	<i>K. Huang, J. Chen, and V. K. N. Lau</i>	3635
Adaptive-Rate Reconstruction of Time-Varying Signals With Application in Compressive Foreground Extraction http://dx.doi.org/10.1109/TSP.2016.2544744	<i>J. F. C. Mota, N. Deligiannis, A. C. Sankaranarayanan, V. Cevher, and M. R. D. Rodrigues</i>	3651
A Functional Composition Approach to Filter Sharpening and Modular Filter Design http://dx.doi.org/10.1109/TSP.2016.2550018	<i>S. Demirtas and A. V. Oppenheim</i>	3667
Mining MOOC Clickstreams: Video-Watching Behavior vs. In-Video Quiz Performance http://dx.doi.org/10.1109/TSP.2016.2546228	<i>C. G. Brinton, S. Buccapatnam, M. Chiang, and H. V. Poor</i>	3677
Multiple Invariance ESPRIT for Nonuniform Linear Arrays: A Coupled Canonical Polyadic Decomposition Approach http://dx.doi.org/10.1109/TSP.2016.2551686	<i>M. Sørensen and L. De Lathauwer</i>	3693
Architectures for Recursive Digital Filters Using Stochastic Computing http://dx.doi.org/10.1109/TSP.2016.2552513	<i>Y. Liu and K. K. Parhi</i>	3705
A Proximal Dual Consensus ADMM Method for Multi-Agent Constrained Optimization http://dx.doi.org/10.1109/TSP.2016.2544743	<i>T.-H. Chang</i>	3719
Optimality of Rate Balancing in Wireless Sensor Networks http://dx.doi.org/10.1109/TSP.2016.2551691	<i>A. Tarighati and J. Jaldén</i>	3735
Stable Support Recovery of Stream of Pulses With Application to Ultrasound Imaging http://dx.doi.org/10.1109/TSP.2016.2552500	<i>T. Bendory, A. Bar-Zion, D. Adam, S. Dekel, and A. Feuer</i>	3750
Successive QCQP Refinement for MIMO Radar Waveform Design Under Practical Constraints http://dx.doi.org/10.1109/TSP.2016.2552501	<i>O. Aldayel, V. Monga, and M. Rangaswamy</i>	3760
Efficient Sampling Set Selection for Bandlimited Graph Signals Using Graph Spectral Proxies http://dx.doi.org/10.1109/TSP.2016.2546233 ..	<i>A. Anis, A. Gadde, and A. Ortega</i>	3775
Sparse Codes Auto-Extractor for Classification: A Joint Embedding and Dictionary Learning Framework for Representation http://dx.doi.org/10.1109/TSP.2016.2550016	<i>Z. Zhang, F. Li, T. W. S. Chow, L. Zhang, and S. Yan</i>	3790
Dynamic Potential Games With Constraints: Fundamentals and Applications in Communications http://dx.doi.org/10.1109/TSP.2016.2551693	<i>S. Zazo, S. V. Macua, M. Sánchez-Fernández, and J. Zazo</i>	3806
correction		
Correction for “On the Convergence of the Iterative Shrinkage/Thresholding Algorithm with a Weakly Convex Penalty” http://dx.doi.org/10.1109/TSP.2016.2572978	<i>I. Bayram</i>	3822



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>



Call for Papers
IEEE Signal Processing Society
IEEE Transactions on Signal and Information Processing over Networks

Special Issue on Distributed Information Processing in Social Networks

Over the past few decades, online social networks such as *Facebook* and *Twitter* have significantly changed the way people communicate and share information with each other. The opinion and behavior of each individual are heavily influenced through interacting with others. These local interactions lead to many interesting collective phenomena such as herding, consensus, and rumor spreading. At the same time, there is always the danger of mob mentality of following crowds, celebrities, or gurus who might provide misleading or even malicious information. Many efforts have been devoted to investigating the collective behavior in the context of various network topologies and the robustness of social networks in the presence of malicious threats. On the other hand, activities in social networks (clicks, searches, transactions, posts, and tweets) generate a massive amount of decentralized data, which is not only big in size but also complex in terms of its structure. Processing these data requires significant advances in accurate mathematical modeling and computationally efficient algorithm design.

Many modern technological systems such as wireless sensor and robot networks are virtually the same as social networks in the sense that the nodes in both networks carry disparate information and communicate with constraints. Thus, investigating social networks will bring insightful principles on the system and algorithmic designs of many engineering networks. An example of such is the implementation of consensus algorithms for coordination and control in robot networks. Additionally, more and more research projects nowadays are data-driven. Social networks are natural sources of massive and diverse big data, which present unique opportunities and challenges to further develop theoretical data processing toolsets and investigate novel applications. This special issue aims to focus on addressing distributed information (signal, data, etc.) processing problems in social networks and also invites submissions from all other related disciplines to present comprehensive and diverse perspectives.

Topics of interest include, but are not limited to:

- Dynamic social networks: time varying network topology, edge weights, etc.
- Social learning, distributed decision-making, estimation, and filtering
- Consensus and coordination in multi-agent networks
- Modeling and inference for information diffusion and rumor spreading
- Multi-layered social networks where social interactions take place at different scales or modalities
- Resource allocation, optimization, and control in multi-agent networks
- Modeling and strategic considerations for malicious behavior in networks
- Social media computing and networking
- Data mining, machine learning, and statistical inference frameworks and algorithms for handling big data from social networks
- Data-driven applications: attribution models for marketing and advertising, trend prediction, recommendation systems, crowdsourcing, etc.
- Other topics associated with social networks: graphical modeling, trust, privacy, engineering applications, etc.

Important Dates:

- Manuscript submission due: September 15, 2016
- First review completed: November 1, 2016
- Revised manuscript due: December 15, 2016
- Second review completed: February 1, 2017
- Final manuscript due: March 15, 2017
- Publication: June 1, 2017

Guest Editors:

- Zhenliang Zhang, Qualcomm Corporate R&D (zhenlian@qti.qualcomm.com)
- Wee Peng Tay, Nanyang Technological University (wptay@ntu.edu.sg)
- Moez Draief, Imperial College London (m.draief@imperial.ac.uk)
- Xiaodong Wang, Columbia University (xw2008@columbia.edu)
- Edwin K. P. Chong, Colorado State University (edwin.chong@colostate.edu)
- Alfred O. Hero III, University of Michigan (hero@eecs.umich.edu)

IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE 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



JUNE 2016

VOLUME 24

NUMBER 6

ITASFA

(ISSN 2329-9290)

regular papers

An Empirical Investigation of Word Class-Based Features for Natural Language Understanding http://dx.doi.org/10.1109/TASLP.2015.2511925	A. Celikyilmaz, R. Sarikaya, M. Jeong, and A. Deoras	994
Feature Adaptation Using Linear Spectro-Temporal Transform for Robust Speech Recognition http://dx.doi.org/10.1109/TASLP.2016.2522646	D. H. H. Nguyen, X. Xiao, E. S. Chng, and H. Li	1006
A Two-Pass Framework of Mispronunciation Detection and Diagnosis for Computer-Aided Pronunciation Training http://dx.doi.org/10.1109/TASLP.2016.2526782	X. Qian, H. Meng, and F. Soong	1020
Text-Independent Phoneme Segmentation Combining EGG and Speech Data http://dx.doi.org/10.1109/TASLP.2016.2533865	L. Chen, X. Mao, and H. Yan	1029
A Framework for Speech Enhancement With Ad Hoc Microphone Arrays http://dx.doi.org/10.1109/TASLP.2016.2537202	V. M. Tavakoli, J. R. Jensen, M. G. Christensen, and J. Benesty	1038

IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING (ISSN 2329-9290) is published bimonthly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), nonmembers \$373.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee of \$31.00 is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2016 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. **Postmaster:** Send address changes to IEEE/ACM Transactions on Audio, Speech, and Language Processing, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.



Candidate Expansion and Prosody Adjustment for Natural Speech Synthesis Using a Small Corpus http://dx.doi.org/10.1109/TASLP.2016.2537982	<i>Y.-Y. Chen, C.-H. Wu, Y.-C. Huang, S.-L. Lin, and J.-F. Wang</i>	1052
A Pairwise Algorithm Using the Deep Stacking Network for Speech Separation and Pitch Estimation http://dx.doi.org/10.1109/TASLP.2016.2540805	<i>X. Zhang, H. Zhang, S. Nie, G. Gao, and W. Liu</i>	1066
An Iterative Approach to Source Counting and Localization Using Two Distant Microphones http://dx.doi.org/10.1109/TASLP.2016.2533859	<i>L. Wang, T.-K. Hon, J. D. Reiss, and A. Cavallaro</i>	1079
A Montage Approach to Sound Texture Synthesis http://dx.doi.org/10.1109/TASLP.2016.2536481	<i>S. O'Leary and A. Röbel</i>	1094
Fast Audio Fingerprinting System Using GPU and a Clustering-Based Technique http://dx.doi.org/10.1109/TASLP.2016.2541303	<i>C. Ouali, P. Dumouchel, and V. Gupta</i>	1106
Using Generic Summarization to Improve Music Information Retrieval Tasks http://dx.doi.org/10.1109/TASLP.2016.2541299	<i>F. Raposo, R. Ribeiro, and D. M. de Matos</i>	1119
Improving Short Utterance Speaker Recognition by Modeling Speech Unit Classes http://dx.doi.org/10.1109/TASLP.2016.2544660	<i>L. Li, D. Wang, C. Zhang, and T. F. Zheng</i>	1129
A Frequency-Domain Adaptive Line Enhancer With Step-Size Control Based on Mutual Information for Harmonic Noise Reduction http://dx.doi.org/10.1109/TASLP.2016.2545920	<i>J. Taghia and R. Martin</i>	1140
EDICS—Editor's Information Classification Scheme http://dx.doi.org/10.1109/TASLP.2016.2567058		1155
Information for Authors http://dx.doi.org/10.1109/TASLP.2016.2567118		1157
announcement		
Call for Papers—Special issue on Sound Scene and Event Analysis http://dx.doi.org/10.1109/TASLP.2016.2567102		1159
Call for Papers—Special Issue on Biosignal-Based Spoken Communication http://dx.doi.org/10.1109/TASLP.2016.2567139		1160

CALL FOR PAPERS

IEEE/ACM Transactions on Audio, Speech and Language Processing
Special issue on **Sound Scene and Event Analysis**

It is still difficult for a machine listening system to demonstrate the same capabilities as human listeners in the analysis of realistic acoustic scenes. Besides speech and music, the analysis of other types of sounds, generally referred to as environmental sounds, is the subject of growing interest from the community and is targeting an ever increasing set of audio categories. In realistic environments, multiple sources are often present simultaneously, and in reverberant conditions, which makes the computational scene analysis challenging.

Typical tasks on audio scene analysis are audio-based scene classification and audio event detection and recognition targeting categories such as “door knocks”, “gunshots”, “crowds”, “car engine noise”, as well as marine mammal and bird species, etc. The wide heterogeneity of possible sounds means that novel types of signal processing and machine learning methods should be developed including novel concepts for audio source segmentation and separation. Beyond recognizing sound scenes and sources of interest, a key task of complex audio scene analysis is sound-source localization.

Further, most of the methods developed until now are probably not tractable on big data so there is also a need for new approaches that are, by design, efficient on large scale problems. Acquiring large scale labelled databases is still problematic and such datasets are most likely collected on heterogeneous sets of acoustic conditions (mobile phone recordings, urban/domestic audio,...) most of which are usually offering a degraded version of the signal of interest with potential variable annotation strategies. Therefore methods to tackle large scale problems also have to be robust against signal degradation, acoustic variability, and annotation variability.

We invite papers on various topics on Sound Scene and Event Analysis, including but not limited to :

- * Audio scene classification;
- * Sound event detection and classification
- * Large-scale environmental audio data sets;
- * Acoustic features for environmental sound analysis;
- * Source localization methods for environmental audio scene analysis
- * Source separation for environmental audio scene analysis
- * Big data in environmental audio;
- * Environmental sound recognition;
- * Computational auditory scene analysis;

The authors are required to follow the Author's Guide for manuscript submission to the IEEE /ACM Transactions on Audio, Speech, and Language Processing at <http://www.signalprocessingsociety.org/publications/periodicals/taslp/>

Important Dates:

Manuscript submission due: July 1st, 2016
First review completed: Sept. 30th 2016
Revised manuscript due: October 20th, 2016
Second review completed: Dec. 1st, 2016
Final manuscript due: Dec. 31st, 2016
Publication date: February 2017

Guest Editors:

Gaël Richard, Télécom ParisTech, France (lead guest editor)
Tuomas Virtanen, Tampere University of Technology, Finland
Juan Pablo Bello, New York University, USA
Nobutaka Ono, National Institute of Informatics, Japan

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



JUNE 2016

VOLUME 25

NUMBER 6

IIPRE4

(ISSN 1057-7149)

PAPERS

Joint Low-Rank and Sparse Principal Feature Coding for Enhanced Robust Representation and Visual Classification http://dx.doi.org/10.1109/TIP.2016.2547180	<i>Z. Zhang, F. Li, M. Zhao, L. Zhang, and S. Yan</i>	2429
A Compressive Multi-Frequency Linear Sampling Method for Underwater Acoustic Imaging http://dx.doi.org/10.1109/TIP.2016.2548243 <i>H. F. Alqadah</i>	2444
Antipodally Invariant Metrics for Fast Regression-Based Super-Resolution http://dx.doi.org/10.1109/TIP.2016.2549362 <i>E. Pérez-Pellitero, J. Salvador, J. Ruiz-Hidalgo, and B. Rosenhahn</i>	2456
Instance-Aware Hashing for Multi-Label Image Retrieval http://dx.doi.org/10.1109/TIP.2016.2545300 <i>H. Lai, P. Yan, X. Shu, Y. Wei, and S. Yan</i>	2469
An Optical Flow-Based Full Reference Video Quality Assessment Algorithm http://dx.doi.org/10.1109/TIP.2016.2548247 <i>M. K. and S. S. Channappayya</i>	2480
Face Aging Effect Simulation Using Hidden Factor Analysis Joint Sparse Representation http://dx.doi.org/10.1109/TIP.2016.2547587 <i>H. Yang, D. Huang, Y. Wang, H. Wang, and Y. Tang</i>	2493
Contrast Driven Elastica for Image Segmentation http://dx.doi.org/10.1109/TIP.2016.2545244 <i>N. Y. El-Zehiry and L. Grady</i>	2508
Fast and Provably Accurate Bilateral Filtering http://dx.doi.org/10.1109/TIP.2016.2548363 <i>K. N. Chaudhury and S. D. Dabhade</i>	2519
Text-Attentional Convolutional Neural Network for Scene Text Detection http://dx.doi.org/10.1109/TIP.2016.2547588 <i>T. He, W. Huang, Y. Qiao, and J. Yao</i>	2529
Cross-View Action Recognition via Transferable Dictionary Learning http://dx.doi.org/10.1109/TIP.2016.2548242 <i>J. Zheng, Z. Jiang, and R. Chellappa</i>	2542
An Empirical Study Into Annotator Agreement, Ground Truth Estimation, and Algorithm Evaluation http://dx.doi.org/10.1109/TIP.2016.2544703 <i>T. A. Lampert, A. Stumpf, and P. Gançarski</i>	2557
Patch-Based Video Denoising With Optical Flow Estimation http://dx.doi.org/10.1109/TIP.2016.2551639 <i>A. Buades, J.-L. Lisani, and M. Miladinović</i>	2573
Modeling, Measuring, and Compensating Color Weak Vision http://dx.doi.org/10.1109/TIP.2016.2539679 <i>S. Oshima, R. Mochizuki, R. Lenz, and J. Chao</i>	2587



A Robust and Simple Measure for Quality-Guided 2D Phase Unwrapping Algorithms http://dx.doi.org/10.1109/TIP.2016.2551370	2601
..... <i>M. Arevalillo-Herrez, F. R. Villatoro, and M. A. Gdeisat</i>	
Sufficient Canonical Correlation Analysis http://dx.doi.org/10.1109/TIP.2016.2551374	2610
..... <i>Y. Guo, X. Ding, C. Liu, and J.-H. Xue</i>	
A Nonlocal TV-Based Variational Method for PolSAR Data Speckle Reduction http://dx.doi.org/10.1109/TIP.2016.2552402	2620
..... <i>X. Nie, H. Qiao, B. Zhang, and X. Huang</i>	
Toner Savings Based on Quasi-Random Sequences and a Perceptual Study for Green Printing http://dx.doi.org/10.1109/TIP.2016.2552641 ..	2635
..... <i>B. Montrucchio and R. Ferrero</i>	
Analyzing the Effect of JPEG Compression on Local Variance of Image Intensity http://dx.doi.org/10.1109/TIP.2016.2553521	2647
..... <i>J. Yang, G. Zhu, and Y.-Q. Shi</i>	
Speeding Up the Bilateral Filter: A Joint Acceleration Way http://dx.doi.org/10.1109/TIP.2016.2549701	2657
..... <i>L. Dai, M. Yuan, and X. Zhang</i>	
Face Recognition With Pose Variations and Misalignment via Orthogonal Procrustes Regression http://dx.doi.org/10.1109/TIP.2016.2551362	2673
..... <i>Y. Tai, J. Yang, Y. Zhang, L. Luo, J. Qian, and Y. Chen</i>	
Lossless Compression of JPEG Coded Photo Collections http://dx.doi.org/10.1109/TIP.2016.2551366	2684
..... <i>H. Wu, X. Sun, J. Yang, W. Zeng, and F. Wu</i>	
A Cognitive Control-Inspired Approach to Object Tracking http://dx.doi.org/10.1109/TIP.2016.2553781	2697
..... <i>A. Mazzu, P. Morerio, L. Marcenaro, and C. S. Regazzoni</i>	
Multi-Label Dictionary Learning for Image Annotation http://dx.doi.org/10.1109/TIP.2016.2549459	2712
..... <i>X.-Y. Jing, F. Wu, Z. Li, R. Hu, and D. Zhang</i>	
Person Re-Identification by Dual-Regularized KISS Metric Learning http://dx.doi.org/10.1109/TIP.2016.2553446	2726
..... <i>D. Tao, Y. Guo, M. Song, Y. Li, Z. Yu, and Y. Y. Tang</i>	
Contrast Sensitivity of the Wavelet, Dual Tree Complex Wavelet, Curvelet, and Steerable Pyramid Transforms http://dx.doi.org/10.1109/TIP.2016.2552725	2739
..... <i>P. Hill, A. Achim, M. E. Al-Mualla, and D. Bull</i>	
Text Detection, Tracking and Recognition in Video: A Comprehensive Survey http://dx.doi.org/10.1109/TIP.2016.2554321	2752
..... <i>X.-C. Yin, Z.-Y. Zuo, S. Tian, and C.-L. Liu</i>	
Double-Tip Artifact Removal From Atomic Force Microscopy Images http://dx.doi.org/10.1109/TIP.2016.2532239	2774
..... <i>Y.-F. Wang, J. I. Kilpatrick, S. P. Jarvis, F. M. Boland, A. Kokaram, and D. Corrigan</i>	
Strokelets: A Learned Multi-Scale Mid-Level Representation for Scene Text Recognition http://dx.doi.org/10.1109/TIP.2016.2555080	2789
..... <i>X. Bai, C. Yao, and W. Liu</i>	
Hermite Snakes With Control of Tangents http://dx.doi.org/10.1109/TIP.2016.2551363	2803
..... <i>V. Uhlmann, J. Fageot, and M. Unser</i>	
Orthographic Perspective Mappings for Consistent Wide-Area Motion Feature Maps From Multiple Cameras http://dx.doi.org/10.1109/TIP.2016.2555079	2817
..... <i>L. O’Gorman and G. Yang</i>	
Convex Sparse Spectral Clustering: Single-View to Multi-View http://dx.doi.org/10.1109/TIP.2016.2553459	2833
..... <i>C. Lu, S. Yan, and Z. Lin</i>	
Compressive Sampling-Based Image Coding for Resource-Deficient Visual Communication http://dx.doi.org/10.1109/TIP.2016.2554320 ..	2844
..... <i>X. Liu, D. Zhai, J. Zhou, X. Zhang, D. Zhao, and W. Gao</i>	
Discriminative Relational Representation Learning for RGB-D Action Recognition http://dx.doi.org/10.1109/TIP.2016.2556940	2856
..... <i>Y. Kong and Y. Fu</i>	
Layer-Based Approach for Image Pair Fusion http://dx.doi.org/10.1109/TIP.2016.2556618	2866
..... <i>C.-H. Son and X.-P. Zhang</i>	
Fusion of Multispectral and Panchromatic Images Based on Morphological Operators http://dx.doi.org/10.1109/TIP.2016.2556944	2882
..... <i>R. Restaino, G. Vivone, M. Dalla Mura, and J. Chanussot</i>	
Image Bit-Depth Enhancement via Maximum A Posteriori Estimation of AC Signal http://dx.doi.org/10.1109/TIP.2016.2553523	2896
..... <i>P. Wan, G. Cheung, D. Florencio, C. Zhang, and O. C. Au</i>	
Sampling Conditions for the Circular Radon Transform http://dx.doi.org/10.1109/TIP.2016.2551364	2910
..... <i>M. Haltmeier</i>	
Two-Dimensional Pattern-Coupled Sparse Bayesian Learning via Generalized Approximate Message Passing http://dx.doi.org/10.1109/TIP.2016.2556582	2920
..... <i>J. Fang, L. Zhang, and H. Li</i>	
<hr/>	
EDICS—Editor’s Information Classification Scheme http://dx.doi.org/10.1109/TIP.2016.2562418	2931
Information for Authors http://dx.doi.org/10.1109/TIP.2016.2562404	2932

IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING



The IEEE Transactions on Computational Imaging publishes 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:



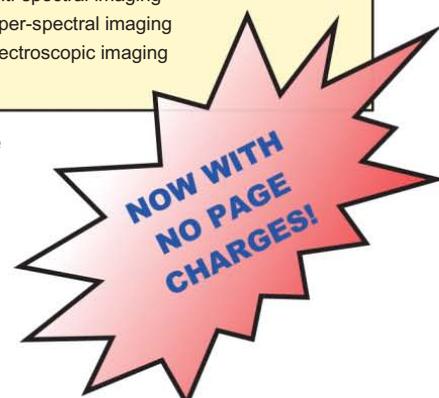
Editor-in-Chief

W. Clem Karl
Boston University

Technical Committee

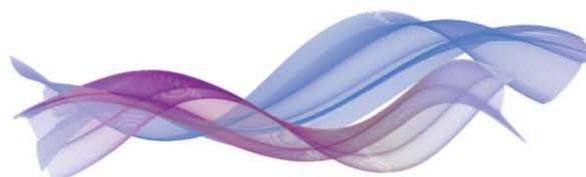
Charles Bouman
Eric Miller
Peter Corcoran
Jong Chul Ye
Dave Brady
William Freeman

<p>Computational Imaging Methods and Models</p> <ul style="list-style-type: none"> • Coded image sensing • Compressed sensing • Sparse and low-rank models • Learning-based models, dictionary methods • Graphical image models • Perceptual models <p>Computational Image Formation</p> <ul style="list-style-type: none"> • Sparsity-based reconstruction • Statistically-based inversion methods • Multi-image and sensor fusion • Optimization-based methods; proximal iterative methods, ADMM <p>Computational Photography</p> <ul style="list-style-type: none"> • Non-classical image capture • Generalized illumination • Time-of-flight imaging • High dynamic range imaging • Plenoptic imaging 	<p>Computational Consumer Imaging</p> <ul style="list-style-type: none"> • Mobile imaging, cell phone imaging • Camera-array systems • Depth cameras, multi-focus imaging • Pervasive imaging, camera networks <p>Computational Acoustic Imaging</p> <ul style="list-style-type: none"> • Multi-static ultrasound imaging • Photo-acoustic imaging • Acoustic tomography <p>Computational Microscopy</p> <ul style="list-style-type: none"> • Holographic microscopy • Quantitative phase imaging • Multi-illumination microscopy • Lensless microscopy • Light field microscopy <p>Imaging Hardware and Software</p> <ul style="list-style-type: none"> • Embedded computing systems • Big data computational imaging • Integrated hardware/digital design 	<p>Tomographic Imaging</p> <ul style="list-style-type: none"> • X-ray CT • PET • SPECT <p>Magnetic Resonance Imaging</p> <ul style="list-style-type: none"> • Diffusion tensor imaging • Fast acquisition <p>Radar Imaging</p> <ul style="list-style-type: none"> • Synthetic aperture imaging • Inverse synthetic aperture imaging <p>Geophysical Imaging</p> <ul style="list-style-type: none"> • Multi-spectral imaging • Ground penetrating radar • Seismic tomography <p>Multi-spectral Imaging</p> <ul style="list-style-type: none"> • Multi-spectral imaging • Hyper-spectral imaging • Spectroscopic imaging
---	--	---



For more information on the IEEE Transactions on Computational Imaging see <http://www.signalprocessingsociety.org/publications/periodicals/tci/>





CAMSAP 2017

Call for Papers

The Seventh IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing

Curaçao, Dutch Antilles

December 10-13, 2017

<http://www.cs.huji.ac.il/conferences/CAMSAP17>



Following the success of the first six editions of the IEEE workshop on Computational Advances in Multi-Sensor Adaptive Processing, we are pleased to announce the seventh workshop in this series. IEEE CAMSAP 2017 will be held in Curaçao, Dutch Antilles, 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	March, 2017
Notification of special session acceptance	March 15, 2015
Submission of papers	July, 2017
Notification of paper acceptance	September, 2017
Final paper submission	



General Chairs

André L. F. de Almeida
andre@gtel.ufc.br
Federal University of Ceará,
Brazil

Martin Haardt
martin.haardt@tu-ilmenau.de
Ilmenau University of Technology,
Germany

Technical Program Chairs

Xiaoli Ma
xiaoli@gatech.edu
Georgia Institute of Technology,
USA

Shahram ShahbazPanahi
shahram.shahbazpanahi@uoiit.ca
University of Ontario Institute of
Technology,
Canada

Finance Chair

Cihan Tepedelenlioglu
cihan@asu.edu
Arizona State University,
USA

Publicity and Publications Chair

Ami Wiesel
ami.wiesel@huji.ac.il
Hebrew University of Jerusalem,
Israel

Local Arrangements Chair

Geert Leus
g.j.t.leus@tudelft.nl
Delft University of Technology,
The Netherlands

Robert W. Heath Jr.
rheath@utexas.edu
University of Texas at Austin,
USA



Join Twitter Conversation
#CAMSAP2017



IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

JUNE 2016

VOLUME 11

NUMBER 6

ITIFA6

(ISSN 1556-6013)

PAPERS

Horizontal and Vertical Side Channel Analysis of a McEliece Cryptosystem http://dx.doi.org/10.1109/TIFS.2015.2509944	1093
..... <i>C. Chen, T. Eisenbarth, I. von Maurich, and R. Steinwandt</i>	
Remanence Decay Side-Channel: The PUF Case http://dx.doi.org/10.1109/TIFS.2015.2512534	1106
..... <i>S. Zeitouni, Y. Oren, C. Wachsmann, P. Koeberl, and A.-R. Sadeghi</i>	
Generating Correlated Digital Certificates: Framework and Applications http://dx.doi.org/10.1109/TIFS.2016.2516818	1117
..... <i>W.-T. Zhu and J. Lin</i>	
Physical Layer Security in Three-Tier Wireless Sensor Networks: A Stochastic Geometry Approach http://dx.doi.org/10.1109/TIFS.2016.2516917	1128
..... <i>Y. Deng, L. Wang, M. ElKashlan, A. Nallanathan, and R. K. Mallik</i>	
Simultaneously Generating Secret and Private Keys in a Cooperative Pairwise-Independent Network http://dx.doi.org/10.1109/TIFS.2016.2516970	1139
..... <i>P. Xu, Z. Ding, X. Dai, and G. K. Karagiannidis</i>	
Against Double Fault Attacks: Injection Effort Model, Space and Time Randomization Based Countermeasures for Reconfigurable Array Architecture http://dx.doi.org/10.1109/TIFS.2016.2518130	1151
..... <i>B. Wang, L. Liu, C. Deng, M. Zhu, S. Yin, and S. Wei</i>	
Identity-Based Proxy-Oriented Data Uploading and Remote Data Integrity Checking in Public Cloud http://dx.doi.org/10.1109/TIFS.2016.2520886	1165
..... <i>H. Wang, D. He, and S. Tang</i>	
Creating Secrets Out of Packet Erasures http://dx.doi.org/10.1109/TIFS.2016.2520887	1177
..... <i>I. Safaka, L. Czap, K. Argyraki, and C. Fragouli</i>	
Specific Emitter Identification via Hilbert–Huang Transform in Single-Hop and Relaying Scenarios http://dx.doi.org/10.1109/TIFS.2016.2520908	1192
..... <i>J. Zhang, F. Wang, O. A. Dobre, and Z. Zhong</i>	
Fingerprint Liveness Detection Using Convolutional Neural Networks http://dx.doi.org/10.1109/TIFS.2016.2520880	1206
..... <i>R. F. Nogueira, R. de Alencar Lotufo, and R. Campos Machado</i>	



Vulnerability Analysis of a Circuit Layout to Hardware Trojan Insertion http://dx.doi.org/10.1109/TIFS.2016.2520910	<i>H. Salmani and M. M. Tehranipoor</i>	1214
One-Class Writer-Independent Offline Signature Verification Using Feature Dissimilarity Thresholding http://dx.doi.org/10.1109/TIFS.2016.2521611	<i>A. Hamadene and Y. Chibani</i>	1226
Imperfect and Perfect Secrecy in Compound Multiple Access Channel With Confidential Message http://dx.doi.org/10.1109/TIFS.2016.2523813	<i>H. Zivari-Fard, B. Akhbari, M. Ahmadian-Attari, and M. R. Aref</i>	1239
ICCDetector: ICC-Based Malware Detection on Android http://dx.doi.org/10.1109/TIFS.2016.2523912	<i>K. Xu, Y. Li, and R. H. Deng</i>	1252
An Efficient File Hierarchy Attribute-Based Encryption Scheme in Cloud Computing http://dx.doi.org/10.1109/TIFS.2016.2523941	<i>S. Wang, J. Zhou, J. K. Liu, J. Yu, J. Chen, and W. Xie</i>	1265
Artificial-Noise-Aided Message Authentication Codes With Information-Theoretic Security http://dx.doi.org/10.1109/TIFS.2016.2524514 ..	<i>X. Wu, Z. Yang, C. Ling, and X.-G. Xia</i>	1278
Design, Evaluation, and Optimization of Physical Unclonable Functions Based on Transient Effect Ring Oscillators http://dx.doi.org/10.1109/TIFS.2016.2524666	<i>A. Cherkaoui, L. Bossuet, and C. Marchand</i>	1291
MagPairing: Pairing Smartphones in Close Proximity Using Magnetometers http://dx.doi.org/10.1109/TIFS.2015.2505626	<i>R. Jin, L. Shi, K. Zeng, A. Pande, and P. Mohapatra</i>	1306
A Dummy-Based Approach for Preserving Source Rate Privacy http://dx.doi.org/10.1109/TIFS.2016.2515050	<i>A. Diyanat, A. Khonsari, and S. P. Shariatpanahi</i>	1321
A Game-Theoretic Framework for Optimum Decision Fusion in the Presence of Byzantines http://dx.doi.org/10.1109/TIFS.2016.2526963 ..	<i>A. Abrardo, M. Barni, K. Kallas, and B. Tondi</i>	1333
Private Cell Retrieval From Data Warehouses http://dx.doi.org/10.1109/TIFS.2016.2527620	<i>X. Yi, R. Paulet, E. Bertino, and G. Xu</i>	1346
Enabling Cloud Storage Auditing With Verifiable Outsourcing of Key Updates http://dx.doi.org/10.1109/TIFS.2016.2528500	<i>J. Yu, K. Ren, and C. Wang</i>	1362
EDICS—Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIFS.2016.2547799		1376
Information for Authors http://dx.doi.org/10.1109/TIFS.2016.2547800		1377

ANNOUNCEMENTS

Call for Papers—IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS http://dx.doi.org/10.1109/TIFS.2016.2559639		1379
Call for Papers—IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING Special Issue on Signal Processing and Machine Learning for Education and Human Learning at Scale http://dx.doi.org/10.1109/TIFS.2016.2559659		1380

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/>

JUNE 2016

VOLUME 18

NUMBER 6

ITMUF8

(ISSN 1520-9210)

PAPERS

3D Audio Signal Processing

Region-Aware 3-D Warping for DIBR <http://dx.doi.org/10.1109/TMM.2016.2539825> J. Jin, A. Wang, Y. Zhao, C. Lin, and B. Zeng 953

Audio/Speech/Language Analysis and Synthesis

Visual Voice Activity Detection in the Wild <http://dx.doi.org/10.1109/TMM.2016.2535357> F. Patrona, A. Iosifidis, A. Tefas, N. Nikolaidis, and I. Pitas 967

Robust Fingertip Detection in a Complex Environment <http://dx.doi.org/10.1109/TMM.2016.2545401> G. Wu and W. Kang 978

Compression and Coding

SSIM-Based Game Theory Approach for Rate-Distortion Optimized Intra Frame CTU-Level Bit Allocation
<http://dx.doi.org/10.1109/TMM.2016.2535254> W. Gao, S. Kwong, Y. Zhou, and H. Yuan 988

Image/Video/Graphics Analysis and Synthesis

CSPS: An Adaptive Pooling Method for Image Classification <http://dx.doi.org/10.1109/TMM.2016.2544099> J. Wang, W. Wang, R. Wang, and W. Gao 1000

Higher-Order Image Co-segmentation <http://dx.doi.org/10.1109/TMM.2016.2545409> W. Wang and J. Shen 1011

Architectures and Design Techniques

A Combined Deblocking Filter and SAO Hardware Architecture for HEVC <http://dx.doi.org/10.1109/TMM.2016.2532606> W. Shen, Y. Fan, Y. Bai, L. Huang, Q. Shang, C. Liu, and X. Zeng 1022

A High-Throughput and Multi-Parallel VLSI Architecture for HEVC Deblocking Filter <http://dx.doi.org/10.1109/TMM.2016.2537217> W. Zhou, J. Zhang, X. Zhou, Z. Liu, and X. Liu 1034



Video Surveillance and Semantic Analysis

Data-Driven Crowd Understanding: A Baseline for a Large-Scale Crowd Dataset <http://dx.doi.org/10.1109/TMM.2016.2542585> C. Zhang, K. Kang, H. Li, X. Wang, R. Xie, and X. Yang 1048

Subjective and Objective Quality Assessment, and User Experience

Multimodal Web Aesthetics Assessment Based on Structural SVM and Multitask Fusion Learning <http://dx.doi.org/10.1109/TMM.2016.2538722> O. Wu, H. Zuo, W. Hu, and B. Li 1062

Binocular Responses for No-Reference 3D Image Quality Assessment <http://dx.doi.org/10.1109/TMM.2016.2542580> ... W. Zhou and L. Yu 1077

Image Sharpness Assessment by Sparse Representation <http://dx.doi.org/10.1109/TMM.2016.2545398> L. Li, D. Wu, J. Wu, H. Li, W. Lin, and A. C. Kot 1085

Saliency-Guided Quality Assessment of Screen Content Images <http://dx.doi.org/10.1109/TMM.2016.2547343> K. Gu, S. Wang, H. Yang, W. Lin, G. Zhai, X. Yang, and W. Zhang 1098

Multimedia Search and Retrieval

Clothing Cosegmentation for Shopping Images With Cluttered Background <http://dx.doi.org/10.1109/TMM.2016.2537783> B. Zhao, X. Wu, Q. Peng, and S. Yan 1111

Multimedia Streaming and Transport

Adaptive Video Streaming With Optimized Bitstream Extraction and PID-Based Quality Control <http://dx.doi.org/10.1109/TMM.2016.2535270> S. Meng, J. Sun, Y. Duan, and Z. Guo 1124

Characterization of Band Codes for Pollution-Resilient Peer-to-Peer Video Streaming <http://dx.doi.org/10.1109/TMM.2016.2535781> A. Fiandrotti, R. Gaeta, and M. Grangetto 1138

Distributed/Cooperative Networks and Communication

On the Optimal Linear Network Coding Design for Information Theoretically Secure Unicast Streaming <http://dx.doi.org/10.1109/TMM.2016.2545403> J. Wang, J. Wang, K. Lu, Y. Qian, and N. Gu 1149

Wireless and Mobile Multimedia

Efficient Cache Placement Strategy in Two-Tier Wireless Content Delivery Network <http://dx.doi.org/10.1109/TMM.2016.2543658> J. Sung, M. Kim, K. Lim, and J.-K. K. Rhee 1163

Multimedia for Immersive Search Space and Personalized Recommendations

Clothes Co-Parsing Via Joint Image Segmentation and Labeling With Application to Clothing Retrieval <http://dx.doi.org/10.1109/TMM.2016.2542983> X. Liang, L. Lin, W. Yang, P. Luo, J. Huang, and S. Yan 1175

Multimedia Social Networks

Estimating Snow Cover From Publicly Available Images <http://dx.doi.org/10.1109/TMM.2016.2535356> R. Fedorov, A. Camerada, P. Fraternali, and M. Tagliasacchi 1187

Multimedia Storytelling and Cross-Modal Translations Between Multimedia Contents

Cross-Modal Correlation Learning by Adaptive Hierarchical Semantic Aggregation <http://dx.doi.org/10.1109/TMM.2016.2535864> Y. Hua, S. Wang, S. Liu, A. Cai, and Q. Huang 1201

Trust In Social Multimedia and Privacy-Protecting Multimedia Analysis

Differentially Private Online Learning for Cloud-Based Video Recommendation With Multimedia Big Data in Social Networks <http://dx.doi.org/10.1109/TMM.2016.2537216> P. Zhou, Y. Zhou, D. Wu, and H. Jin 1217

ANNOUNCEMENT

Introducing IEEE Collabratec <http://dx.doi.org/10.1109/TMM.2016.2563579> 1230

Information for Authors <http://dx.doi.org/10.1109/TMM.2016.2563580> 1231

IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING


www.ieee.org/sp/index.html

JUNE 2016

VOLUME 10

NUMBER 4

IJSTGY

(ISSN 1932-4553)

ISSUE ON STRUCTURED MATRICES IN SIGNAL AND DATA PROCESSING

EDITORIAL

Introduction to the Issue on Structured Matrices in Signal and Data Processing <http://dx.doi.org/10.1109/JSTSP.2016.2553398>
 *M. B. Wakin, R. Gribonval, V. Koivunen, J. Romberg, and J. Wright* 605

OVERVIEW ARTICLE

An Overview of Low-Rank Matrix Recovery From Incomplete Observations <http://dx.doi.org/10.1109/JSTSP.2016.2539100>
 *M. A. Davenport and J. Romberg* 608

PAPERS

A Characterization of Deterministic Sampling Patterns for Low-Rank Matrix Completion <http://dx.doi.org/10.1109/JSTSP.2016.2537145> ...
 *D. L. Pimentel-Alarcón, N. Boston, and R. D. Nowak* 623

Hankel Low-Rank Matrix Completion: Performance of the Nuclear Norm Relaxation <http://dx.doi.org/10.1109/JSTSP.2016.2535182>
 *K. Usevich and P. Comon* 637

Structured Low-Rank Matrix Factorization for Haplotype Assembly <http://dx.doi.org/10.1109/JSTSP.2016.2547860>
 *C. Cai, S. Sanghavi, and H. Vikalo* 647

Off-the-Grid Low-Rank Matrix Recovery and Seismic Data Reconstruction <http://dx.doi.org/10.1109/JSTSP.2016.2555482>
 *O. López, R. Kumar, Ö. Yılmaz, and F. J. Herrmann* 658

Beyond Low Rank + Sparse: Multiscale Low Rank Matrix Decomposition <http://dx.doi.org/10.1109/JSTSP.2016.2545518>
 *F. Ong and M. Lustig* 672



Flexible Multilayer Sparse Approximations of Matrices and Applications http://dx.doi.org/10.1109/JSTSP.2016.2543461	688
..... <i>L. L. Magoarou, and R. Gribonval</i>	
Generic Uniqueness of a Structured Matrix Factorization and Applications in Blind Source Separation http://dx.doi.org/10.1109/JSTSP.2016.2526971	701
..... <i>I. Domanov and L. D. Lathauwer</i>	
A Provably Efficient Algorithm for Separable Topic Discovery http://dx.doi.org/10.1109/JSTSP.2016.2555240	712
..... <i>W. Ding, P. Ishwar, and V. Saligrama</i>	
Linearized Kernel Dictionary Learning http://dx.doi.org/10.1109/JSTSP.2016.2555241	726
..... <i>A. Golts and M. Elad</i>	
Fast Robust PCA on Graphs http://dx.doi.org/10.1109/JSTSP.2016.2555239	740
..... <i>N. Shahid, N. Perraudin, V. Kalofolias, G. Puy, and P. Vandergheynst</i>	
Tensor CP Decomposition With Structured Factor Matrices: Algorithms and Performance http://dx.doi.org/10.1109/JSTSP.2016.2509907 ...	757
..... <i>J. H. de M. Goulart, M. Boizard, R. Boyer, G. Favier, and P. Comon</i>	
STFT Phase Retrieval: Uniqueness Guarantees and Recovery Algorithms http://dx.doi.org/10.1109/JSTSP.2016.2549507	770
..... <i>K. Jaganathan, Y. C. Eldar, and B. Hassibi</i>	
Guaranteed Blind Sparse Spikes Deconvolution via Lifting and Convex Optimization http://dx.doi.org/10.1109/JSTSP.2016.2543462	782
..... <i>Y. Chi</i>	
Parametric Bilinear Generalized Approximate Message Passing http://dx.doi.org/10.1109/JSTSP.2016.2539123	795
..... <i>J. T. Parker and P. Schniter</i>	
Learning-Based Compressive Subsampling http://dx.doi.org/10.1109/JSTSP.2016.2548442	809
..... <i>L. Baldassarre, Y.-H. Li, J. Scarlett, B. Gözcü, I. Bogunovic, and V. Cevher</i>	
Information for Authors http://dx.doi.org/10.1109/JSTSP.2016.2560786	823

Call for Papers

IEEE Journal of Selected Topics in Signal Processing

Special Issue on Spoofing and Countermeasures for Automatic Speaker Verification

Automatic speaker verification (ASV) offers a low-cost and flexible biometric solution to person authentication. While the reliability of ASV systems is now considered sufficient to support mass-market adoption, there are concerns that the technology is vulnerable to spoofing, also referred to as presentation attacks. Replayed, synthesized and converted speech spoofing attacks can all project convincing, high-quality speech signals that are representative of other, specific speakers and thus present a genuine threat to the reliability of ASV systems.

Recent years have witnessed a movement in the community to develop spoofing countermeasures, or presentation attack detection (PAD) technology to help protect ASV systems from fraud. These efforts culminated in the first standard evaluation platform for the assessment of spoofing and countermeasures of automatic speaker verification – the Automatic Speaker Verification Spoofing and Countermeasures Challenge (ASVspoof) – which was held as a special session at Interspeech 2015.

This special issue is expected to present original papers describing the very latest developments in spoofing and countermeasures for ASV. The focus of the special issue includes, but is not limited to the following topics related to spoofing and countermeasures for ASV:

- vulnerability analysis of previously unconsidered spoofing methods;
- advanced methods for standalone countermeasures;
- advanced methods for joint ASV and countermeasure modelling;
- information theoretic approaches for the assessment of spoofing and countermeasures;
- spoofing and countermeasures in adverse acoustic and channel conditions;
- generalized and speaker-dependent countermeasures;
- speaker obfuscation, impersonation, de-identification, disguise, evasion and adapted countermeasures;
- analysis and comparison of human performance in the face of spoofing;
- new evaluation protocols, datasets, and performance metrics for the assessment of spoofing and countermeasures for ASV;
- countermeasure methods using other modality or multimodality that are applicable to speaker verification

Also invited are submissions of exceptional quality with a tutorial or overview nature. Creative papers outside the areas listed above but related to the overall scope of the special issue are also welcome. Prospective authors can contact the Guest Editors to ascertain interest on such topics.

Prospective authors should visit <http://www.signalprocessingsociety.org/publications/periodicals/jstsp/> for submission information. Manuscripts should be submitted at <http://mc.manuscriptcentral.com/jstsp-ieee> and will be peer reviewed according to standard IEEE processes.

Important Dates:

- Manuscript submission due: August 1, 2016
- First review completed: October 15, 2016
- Revised manuscript due: December 1, 2016
- Second review completed: February 1, 2017
- Final manuscript due: March 1, 2017
- Publication date: June, 2017

Guest Editors:

Junichi Yamagishi, National Institute of Informatics, Japan, email: jyamagis@nii.ac.jp

Nicholas Evans, EURECOM, France, email: evans@eurecom.fr

Tomi Kinnunen, University of Eastern Finland, Finland, email: tomi.kinnunen@uef.fi

Phillip L. De Leon, New Mexico State University & VoiceCipher, USA, email: pdeleon@nmsu.edu

Isabel Trancoso, INESC-ID, Portugal, email: Isabel.Trancoso@inesc-id.pt

CALL FOR PAPERS
IEEE Journal of Selected Topics in Signal Processing
Special Issue on Signal Processing and Machine Learning
for Education and Human Learning at Scale

Aims and Scope

The surge in popularity of Massive Open Online Courses (MOOCs) and other online and blended learning platforms has demonstrated the potential of the Internet for scaling education. While advances in technology have enabled content delivery to massive numbers of students, these platforms remain limited in their ability to provide an effective learning experience for each individual.

Recent advances in machine learning and signal processing offer promising avenues to move beyond this “one size fits all” educational approach. The key is that today’s learning technology platforms can capture big data about learners as they proceed through courses. Examples of learning data include performance on homeworks and exams, click actions made while watching lecture videos or interacting with simulations, the social learning networks formed among the students, and the content posted on discussion forums. Going even further, prototype platforms are being built that use cameras and other sensors to continuously monitor students’ affect and engagement. The large volumes of empirical learning data being collected present novel opportunities to study the process of student learning, to design systems that improve learning at scale by closing the learning feedback loop.

This special issue of IEEE J-STSP will showcase the research from the signal processing community that is providing leadership in advancing effective learning at scale. Particularly of interest to this special issue will be novel methods for defining and extracting signals of a student’s behavior and performance from big learning data and using these measures in the design of intelligent algorithms and systems.

Topics of interest in the special issue include (but are not limited to):

- Processing and Representing Learning Behavioral Data
- Generative/Low Dimensional Modeling of Student Learning
- Learner Knowledge Tracing and Performance Prediction
- Social Learning Networks
- Algorithms for Identifying Learner Collaborations
- Automating Course/Content Individualization, Automatic Grading Methods, Automatic Feedback Generation
- Learning Analytics with Actionable Intelligence for Instructors
- Algorithms for More Effective Peer Grading Allocation
- Relationships between Learning Behavior, Performance, and Content
- Machine Vision Algorithms for Processing Student Biometric Data
- Trials for Demonstrating Efficacy for Learners and/or Instructors

Important Dates:

Manuscript submission due: October 1, 2016

First review completed: December 15, 2016

Revised manuscript due: February 1, 2017

Second review completed: May 1, 2017

Final manuscript due: June 15, 2017

Publication: August 2017

Prospective authors should visit <http://www.signalprocessingsociety.org/publications/periodicals/jstsp/> for information on paper submission. Manuscripts should be submitted using the Manuscript Central system at <http://mc.manuscriptcentral.com/jstsp-ieee>.

Guest Editors:

Mihaela van der Schaar (Lead Guest Editor), University of California Los Angeles, USA, mihaela@ee.ucla.edu

Mung Chiang, Princeton University, USA, chiangm@princeton.edu

Richard Baraniuk, Rice University, USA, richb@rice.edu

Jonathan Chung-Kuan Huang, Google, USA, jonathanhuang@google.com

Shengdong Zhao, National University of Singapore, Singapore, zhaosd@comp.nus.edu.sg

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY


www.ieee.org/sp/index.html

JUNE 2016

VOLUME 23

NUMBER 6

ISPLEM

(ISSN 1070-9908)

LETTERS

Visual Face Recognition Using Bag of Dense Derivative Depth Patterns http://dx.doi.org/10.1109/LSP.2016.2553784	771
..... <i>T. Mantecón, C. R. del-Blanco, F. Jaureguizar, and N. García</i>	
Fast Frequency Estimation Algorithm by Least Squares Phase Unwrapping http://dx.doi.org/10.1109/LSP.2016.2555933	776
..... <i>Z. Xu, T. Lu, and B. Huang</i>	
A Subband-Based Stationary-Component Suppression Method Using Harmonics and Power Ratio for Reverberant Speech Recognition http://dx.doi.org/10.1109/LSP.2016.2554888	780
..... <i>B. J. Cho, H. Kwon, J.-W. Cho, C. Kim, R. M. Stern, and H.-M. Park</i>	
Multivideo Object Cosegmentation for Irrelevant Frames Involved Videos http://dx.doi.org/10.1109/LSP.2016.2557346	785
..... <i>J. Zhang, K. Li, and W. Tao</i>	
Maximum-Likelihood Blind Synchronization for GFDM Systems http://dx.doi.org/10.1109/LSP.2016.2552234	790
..... <i>P.-S. Wang and D. W. Lin</i>	
Semisupervised Multilabel Learning With Joint Dimensionality Reduction http://dx.doi.org/10.1109/LSP.2016.2554361	795
..... <i>T. Yu and W. Zhang</i>	
1-D Filter for Ring Artifact Suppression http://dx.doi.org/10.1109/LSP.2016.2554363	800
..... <i>V. Titarenko</i>	
ECG Authentication System Design Based on Signal Analysis in Mobile and Wearable Devices http://dx.doi.org/10.1109/LSP.2016.2531996	805
..... <i>S. J. Kang, S. Y. Lee, H. I. Cho, and H. Park</i>	
Bias-Compensated Normalized Subband Adaptive Filter Algorithm http://dx.doi.org/10.1109/LSP.2016.2554886	809
..... <i>Z. Zheng and H. Zhao</i>	
Pseudo-Pilot: A Novel Paradigm of Channel Estimation http://dx.doi.org/10.1109/LSP.2016.2557347	814
..... <i>Y. Ma</i>	
Saliency Detection for Stereoscopic Images Based on Depth Confidence Analysis and Multiple Cues Fusion http://dx.doi.org/10.1109/LSP.2016.2557347	819
..... <i>R. Cong, J. Lei, C. Zhang, Q. Huang, X. Cao, and C. Hou</i>	
Improving the Goertzel–Blahut Algorithm http://dx.doi.org/10.1109/LSP.2016.2557486	824
..... <i>S. V. Fedorenko</i>	
Cooperative Localization for Mobile Networks: A Distributed Belief Propagation – Mean Field Message Passing Algorithm http://dx.doi.org/10.1109/LSP.2016.2550534 ...	828
..... <i>B. Çakmak, D. N. Urup, F. Meyer, T. Pedersen, B. H. Fleury, and F. Hlawatsch</i>	

IEEE SIGNAL PROCESSING LETTERS (ISSN 1070–9908) is published quarterly in print and monthly online by the Institute of Electrical and Electronics Engineers, Inc. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society/Council, or its members. **IEEE Corporate Office:** 3 Park Avenue, 17th Floor, New York, NY 10016-5997. **IEEE Operations Center:** 445 Hoes Lane, Piscataway, NJ 08854-4141. **NJ Telephone:** +1 732 981 0060. **Price/Publication Information:** Individual copies: IEEE Members \$20.00 (first copy only), nonmembers \$340.00 per copy. (Note: Postage and handling charge not included.) Member and nonmember subscription prices available upon request. Available in microfiche and microfilm. **Copyright and Reprint Permissions:** Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the per-copy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For all other copying, reprint, or republication permission, write to Copyrights and Permissions Department, IEEE Publications Administration, 445 Hoes Lane, Piscataway, NJ 08854-4141. Copyright © 2016 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. **Postmaster:** Send address changes to IEEE Signal Processing Letters, IEEE, 445 Hoes Lane, Piscataway, NJ 08854-4141. GST Registration No. 125634188. CPC Sales Agreement #40013087. Return undeliverable Canada addresses to: Pitney Bowes IMEX, P.O. Box 4332, Stanton Rd., Toronto, ON M5W 3J4, Canada. IEEE prohibits discrimination, harassment and bullying. For more information visit <http://www.ieee.org/nondiscrimination>. Printed in U.S.A.

Rotation Invariant Texture Description Using Symmetric Dense Microblock Difference http://dx.doi.org/10.1109/LSP.2016.2561311	833
..... <i>R. Mehta and K. Egiazarian</i>	
Saliency Detection Via Similar Image Retrieval http://dx.doi.org/10.1109/LSP.2016.2558489	838
..... <i>L. Ye, Z. Liu, X. Zhou, L. Shen, and J. Zhang</i>	
Generalized Brillinger-Like Transforms http://dx.doi.org/10.1109/LSP.2016.2556714	843
..... <i>A. Torokhti and P. Soto-Quiros</i>	
New Results on Generalized Fractional Programming Problems With Toeplitz Quadratics http://dx.doi.org/10.1109/LSP.2016.2555880	848
..... <i>A. Aubry, V. Carotenuto, and A. D. Maio</i>	
Image Dehazing Using Quadtree Decomposition and Entropy-Based Contextual Regularization http://dx.doi.org/10.1109/LSP.2016.2559805	853
..... <i>N. Baig, M. M. Riaz, A. Ghafoor, and A. M. Siddiqui</i>	
Geometric-Algebra LMS Adaptive Filter and Its Application to Rotation Estimation http://dx.doi.org/10.1109/LSP.2016.2558461	858
..... <i>W. B. Lopes, A. Al-Nuaimi, and C. G. Lopes</i>	
Scalable Multisensor Multitarget Tracking Using the Marginalized δ -GLMB Density http://dx.doi.org/10.1109/LSP.2016.2557078	863
..... <i>C. Fantacci and F. Papi</i>	
Efficient Demodulation of General APSK Constellations http://dx.doi.org/10.1109/LSP.2016.2560241	868
..... <i>M. Sandell, F. Tosato, and A. Ismail</i>	
Statistical Performance Analysis of the Adaptive Orthogonal Rejection Detector http://dx.doi.org/10.1109/LSP.2016.2550495	873
..... <i>W. Liu, J. Liu, X. Hu, Z. Tang, L. Huang, and Y.-L. Wang</i>	
An Inner-Product Calculus for Periodic Functions and Curves http://dx.doi.org/10.1109/LSP.2016.2555139	878
..... <i>A. Badoual, D. Schmitter, and M. Unser</i>	
Transceiver Design to Maximize the Weighted Sum Secrecy Rate in Full-Duplex SWIPT Systems http://dx.doi.org/10.1109/LSP.2016.2553171	883
..... <i>Y. Wang, R. Sun, and X. Wang</i>	
Adaptive Part-Level Model Knowledge Transfer for Gender Classification http://dx.doi.org/10.1109/LSP.2016.2555480	888
..... <i>Y. Gao, Z. Li, and Y. Qiao</i>	
Robust Cross-view Hashing for Multimedia Retrieval http://dx.doi.org/10.1109/LSP.2016.2517093	893
..... <i>X. Shen, F. Shen, Q.-S. Sun, Y.-H. Yuan, and H. T. Shen</i>	

EDICS—Editors' Information Classification Scheme http://dx.doi.org/10.1109/LSP.2016.2567701	898
Information for Authors http://dx.doi.org/10.1109/LSP.2016.2567702	899

IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING

A PUBLICATION OF
IEEE SIGNAL PROCESSING SOCIETY
IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY
IEEE CONSUMER ELECTRONICS SOCIETY



TECHNICALLY CO-SPONSORED BY
IEEE GEOSCIENCE AND REMOTE SENSING SOCIETY



JUNE 2016

VOLUME 2

NUMBER 02

ITCIAJ

(ISSN 2333-9403)

PAPERS

Computational Imaging Methods and Models

Locally Similar Sparsity-Based Hyperspectral Compressive Sensing Using Unmixing http://dx.doi.org/10.1109/TCLI.2016.2542002	86
..... L. Zhang, W. Wei, Y. Zhang, H. Yan, F. Li, and C. Tian	
A Geometric Model to Characterize Annihilation Positions Associated With Scattered Coincidences in PET: A Simulation-Based Study http://dx.doi.org/10.1109/TCLI.2016.2549742	101
..... H. Sun and S. Pistorius	
Video Super-Resolution With Convolutional Neural Networks http://dx.doi.org/10.1109/TCLI.2016.2532323	109
..... A. Kappeler, S. Yoo, Q. Dai, and A. K. Katsaggelos	
Models of Monocular and Binocular Visual Perception in Quality Assessment of Stereoscopic Images http://dx.doi.org/10.1109/TCLI.2016.2538720	123
..... F. Shao, W. Lin, G. Jiang, and Q. Dai	
<i>Computational Image Formation</i>	
Joint Reconstruction of Absorbed Optical Energy Density and Sound Speed Distributions in Photoacoustic Computed Tomography: A Numerical Investigation http://dx.doi.org/10.1109/TCLI.2016.2523427	136
..... C. Huang, K. Wang, R. W. Schoonover, L. V. Wang, and M. A. Anastasio	
<i>Computational Imaging Systems</i>	
Blind X-Ray CT Image Reconstruction From Polychromatic Poisson Measurements http://dx.doi.org/10.1109/TCLI.2016.2523431	150
..... R. Gu and A. Dogandžić	
Numerical Inversion of a Broken Ray Transform Arising in Single Scattering Optical Tomography http://dx.doi.org/10.1109/TCLI.2016.2531581	166
..... G. Ambartsoumian and S. Roy	

EDICS—Editor’s Classification Information Scheme http://dx.doi.org/10.1109/TCLI.2016.2560539	174
Information for Authors http://dx.doi.org/10.1109/TCLI.2016.2560540	175



IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS

A PUBLICATION OF
THE IEEE SIGNAL PROCESSING SOCIETY
THE IEEE COMMUNICATIONS SOCIETY
THE IEEE COMPUTER SOCIETY



JUNE 2016

VOLUME 2

NUMBER 2

ITSIBW

(ISSN 2373-776X)

PAPERS

Adaptation, Detection, Estimation, and Learning

- Distributed Detection With Vector Quantizer <http://dx.doi.org/10.1109/TSIPN.2016.2524572> W. Zhao and L. Lai 105
- NEXT: In-Network Nonconvex Optimization <http://dx.doi.org/10.1109/TSIPN.2016.2524588> P. D. Lorenzo and G. Scutari 120
- Graph Signal Denoising via Trilateral Filter on Graph Spectral Domain <http://dx.doi.org/10.1109/TSIPN.2016.2532464> M. Onuki, S. Ono, M. Yamagishi, and Y. Tanaka 137

Communication, Networking, and Sensing

- Energy-Assisted Information Detection for Simultaneous Wireless Information and Power Transfer: Performance Analysis and Case Studies <http://dx.doi.org/10.1109/TSIPN.2016.2539682> C.-H. Chang, R. Y. Chang, and F.-T. Chien 149

Data Analysis

- Principal Patterns on Graphs: Discovering Coherent Structures in Datasets <http://dx.doi.org/10.1109/TSIPN.2016.2524500> K. Benzi, B. Ricaud, and P. Vandergheynst 160
- Location of Things: Geospatial Tagging for IoT Using Time-of-Arrival <http://dx.doi.org/10.1109/TSIPN.2016.2531422> I. Nevat, G. W. Peters, K. Avnit, F. Septier, and L. Clavier 174

Emerging Topics and Applications

- Node Dominance: Revealing Community and Core-Periphery Structure in Social Networks <http://dx.doi.org/10.1109/TSIPN.2016.2527923> .. J. Gamble, H. Chintakunta, A. Wilkerson, and H. Krim 186
- Data Denoising and Compression for Smart Grid Communication <http://dx.doi.org/10.1109/TSIPN.2016.2539680> J. Khan, S. Bhuiyan, G. Murphy, and J. Williams 200

Modeling and Analysis

- Extraction of Temporal Network Structures From Graph-Based Signals <http://dx.doi.org/10.1109/TSIPN.2016.2530562> R. Hamon, P. Borgnat, P. Flandrin, and C. Robardet 215

- EDICS—Editor's Information Classification Scheme <http://dx.doi.org/10.1109/TSIPN.2016.2558720> 227
- Information for Authors <http://dx.doi.org/10.1109/TSIPN.2016.2558721> 228





Sponsored by

(Pending)



Supported by

**General Chairs****Christos Douligeris***University of Piraeus, Greece***Andreas Pitsillides***University of Cyprus, Cyprus***Andreas Spanias***Arizona State University, USA***Technical Program Chairs****Ioannis Kyriakides***University of Nicosia, Cyprus***Veselin Rakocevic***City University, UK***Thanos Stouraitis***Khalifa University, UAE***Registration & Finance Chair****Reda Ammar***University of Connecticut, USA***Publication Co-Chairs****Josephine Antoniou***University of Central Lancashire, Cyprus***Christophoros Christophorou***University of Cyprus, Cyprus***Angeliki Tsioliaridou***Foundation of Research and Technology, Greece***Tutorials Co-Chairs****Marios Lestas***Frederick University, Cyprus***Vasos Vassiliou***University of Cyprus, Cyprus***Plenary & Special Sessions Co-Chairs****Vicky Lesta-Papadopoulou***European University, Cyprus***Christos Liaskos***Foundation of Research and Technology, Greece***Publicity Co-Chairs****Suvendi Rimer***University of Johannesburg, South Africa***Cristiano Silva***Universidade Federal de São João Del-rei***Local Arrangements Chair****George A. Papadopoulos***University of Cyprus, Cyprus***Web Manager****Kyriakos Georgiades***Easy Conferences, Cyprus***Steering Committee****Esam Abdel-Raheem***University of Windsor, Canada***Reda Ammar***University of Connecticut, USA***Fayez Elguibaly***University of Victoria, Canada***Adel Elmaghraby***University of Louisville, USA (Chair)***Ahmed Tantawy***IBM, Egypt / USA***Ahmed Tewfik***University of Minnesota, USA***Call for Papers**

IEEE ISSPIT 2016 is the sixteenth in a series of international symposia that aims to cover a wide range of topics in the intersection of signal processing and information technology and to become a fertile ground for discussions between the two research and development communities. Apart from sessions that will present new research results, tutorials and special sessions will be offered as well. Papers describing original work are invited in the general fields covered by ISSPIT, with an emphasis on the topics listed below. Accepted papers will be published in the *Proceedings of IEEE ISSPIT 2016* and will be available via *IEEE Xplore*. Acceptance will be based on quality, relevance, and originality. A contest for Best Paper Awards (senior and student) will be held.

Papers are invited in the following (not exclusive) **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
- Big Data Analytics in Signal Processing and IT
- Radar Signal Processing
- Engineering Systems-of-Systems
- Internet Software Architectures
- Multimedia and Image-Based Systems
- Computer Networks
- Mobile Information Systems
- E-Commerce
- Bioinformatics and Bioengineering
- Information Processing
- Geographical Information Systems
- Object-Based Software Engineering
- Speech Processing
- Neural Networks
- Social Networks Analysis
- Internet of Things

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. More information regarding submissions can be found on the ISSPIT website: www.cyprusconferences.org/isspit2016. 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	Aug. 5, 2016
Regular paper submission	Sep. 15, 2016
Notification of acceptance	Oct. 20, 2016
Final version paper with registration	Nov. 10, 2016

Academic Inquiries

Prof. George A. Papadopoulos
george@cs.ucy.ac.cy

Questions including Accommodation and Travel Arrangements

Easy Conferences
info@easyconferences.eu

www.cyprusconferences.org/isspit2016

IEEE Signal Processing MAGAZINE

Volume 33 | Number 4 | July 2016

MOOC ADVENTURES IN SIGNAL PROCESSING

Acoustic Microphone
Geometry Calibration

Mobile Image Processing
for Health Monitoring

Continuous User Authentication

Application-Oriented Hands-On
Activities for SP Courses

IEEE
Signal Processing Society

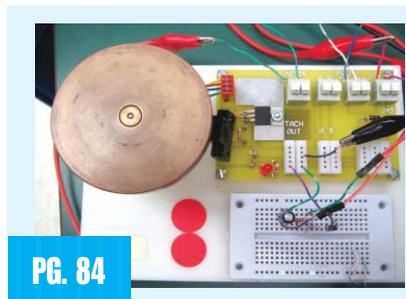
IEEE

Contents

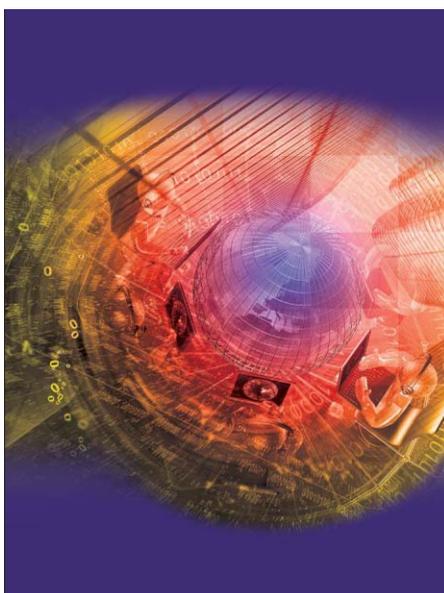
Volume 33 | Number 4 | July 2016

FEATURES

- 14 ACOUSTIC MICROPHONE GEOMETRY CALIBRATION**
Axel Plinge, Florian Jacob, Reinhold Haeb-Umbach, and Gernot A. Fink
- 30 SMARTPHONE AND MOBILE IMAGE PROCESSING FOR ASSISTED LIVING**
Hossein Nejati, Victor Pomponiu, Thanh-Toan Do, Yiren Zhou, Sahar Irvani, and Ngai-Man Cheung
- 49 CONTINUOUS USER AUTHENTICATION ON MOBILE DEVICES**
Vishal M. Patel, Rama Chellappa, Deepak Chandra, and Brandon Barbello



PG. 84



ON THE COVER

The last four years have witnessed an explosion of interest and activity in massive open online courses (MOOCs). Online teaching has caused much debate, mostly because of its undeniably disruptive nature with respect to the standard college education paradigm. Check out the article on page 62, as the authors share their collective experiences and insights learned through the design, development, delivery, and management of online DSP courses.

COVER IMAGE: ©ISTOCKPHOTO.COM/ARCHERIX

62 MOOC ADVENTURES IN SIGNAL PROCESSING

Thomas A. Baran,
Richard G. Baraniuk,
Alan V. Oppenheim,
Paolo Prandoni, and
Martin Vetterli

COLUMNS

- 6 Society News**
New Society Officer Elected for 2017 and Nominations Open for 2016 SPS Awards
- 8 Reader's Choice**
Top Downloads in IEEE Xplore
- 10 Special Reports**
Signal Processing Plays a Key Role in Wireless Research
John Edwards
- 84 SP Education**
Lessons Learned from Implementing Application-Oriented Hands-On Activities for Continuous-Time Signal Processing Courses
Mario Simoni and Maurice Aburdene
- 90 Life Sciences**
Understanding and Predicting Epilepsy
Christophe Bernard
- 96 Tips & Tricks**
A General Design Method for FIR Compensation Filters in $\Delta\Sigma$ ADCs
Zhe Chen, Shuwen Wang, and Fuliang Yin



PG. 90

IEEE SIGNAL PROCESSING MAGAZINE (ISSN 1053-5888) (ISPREG) is published bimonthly by the Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, 17th Floor, New York, NY 10016-5997 USA (+1 212 419 7900). Responsibility for the contents rests upon the authors and not the IEEE, the Society, or its members. Annual member subscriptions included in Society fee. Nonmember subscriptions available upon request. **Individual copies:** IEEE Members US\$20.00 (first copy only), nonmembers US\$213.00 per copy. Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limits of U.S. Copyright Law for private use of patrons: 1) those post-1977 articles that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923 USA; 2) pre-1978 articles without fee. Instructors are permitted to photocopy isolated articles for noncommercial classroom use without fee. **For all other copying, reprint, or republication permission,** write to IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854 USA. Copyright © 2016 by the Institute of Electrical and Electronics Engineers, Inc. All rights reserved. Periodicals postage paid at New York, NY, and at additional mailing offices. **Postmaster:** Send address changes to IEEE Signal Processing Magazine, IEEE, 445 Hoes Lane, Piscataway, NJ 08854 USA. Canadian GST #125634188 **Printed in the U.S.A.**

Digital Object Identifier 10.1109/MSP.2016.2547159

Organizing Committee

Co-Chairs

Najim Dehak
Johns Hopkins University

Pedro Torres-Carrasquillo
MIT Lincoln Laboratory

Technical Co-Chairs

Dilek Hakkani-Tur
Microsoft

Julia Hirschberg
Columbia University

Douglas Reynolds
MIT Lincoln Laboratory

Frank Seide
Microsoft

Zheng Hua Tan
Aalborg University

Dan Povey
Johns Hopkins University

Regional Publicity Chairs

Nancy Chen
I2R, A*STAR

Reinhold Häb-Umbach
Paderborn University

Joseph P. Campbell, Jr.
MIT Lincoln Laboratory

Finance Chair

Douglas O'Shaughnessy
INRS

Sponsorship Chairs

Geoff Zweig
Microsoft

Erik McDermott
Google

Jerome Bellegarda
Apple

Publication Chairs

Patrick Cardinal
ETS

John H. L. Hansen
UT Dallas

Tutorial, Panel and Invited Speakers

Roger K. Moore
Sheffield University

Tara Sainath
Google

Demo Chairs

Chuck Wooters
AST

Andreas Stocke
Microsoft

Advisory Board

James R. Glass
MIT

Jason Williams
Microsoft

2016 IEEE Workshop on Spoken Language Technology



13-16 December 2016 • San Juan, Puerto Rico

The Sixth IEEE Workshop on Spoken Language Technology (SLT) will be held from December 13-16, 2016 in San Juan, Puerto Rico. The theme for this year will be "machine learning from signal to concepts". The workshop is expected to provide researchers around the world the opportunity to interact and present their newest and most advanced research in the fields of speech and language processing. The program for SLT 2016 will include oral and posters sessions, keynotes, plus invited speakers in the field of spoken language as well as tutorials and multiple special sessions.

Topics

Submission of papers is desired on a large variety of areas of spoken language technology, with emphasis on the following topics on previous workshops:

Speech recognition and synthesis	Human/computer interaction
Spoken language understanding	Spoken dialog systems
Spoken document retrieval	Speech data mining
Question answering from speech	Spoken document summarization
Assistive technologies	Spoken language databases
Natural language processing	Speaker/language recognition
Educational and healthcare applications	Multimodal processing

Venue

IEEE SLT 2016 will take place in San Juan, Puerto Rico at the InterContinental Hotel in the tourist area of Isla Verde. These areas feature beautiful beaches and a vibrant night life besides a large number of dining options. Additionally, the Old San Juan area is just a few miles away. Additional details about SLT 2016 can be found at: www.slt2016.org

Important Dates

Special Session Proposals:	June 8, 2016
Paper Submission:	July 22, 2016
Notification of Review Results:	September 14, 2016
Demo Submission:	September 16, 2016
Early Registration Deadline:	October 14, 2016
Workshop:	December 13-16, 2016

Submission Details

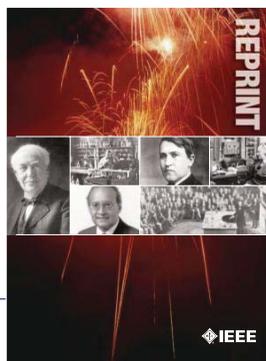
Authors are invited to prepare a full-length manuscript of 4-6 pages, including reference materials and figures, to the SLT 2016 website: www.slt2016.org



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.

Authorized Signature: _____ Date: _____



2016 IEEE MEMBERSHIP APPLICATION

(students and graduate students must apply online)

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 | 9. Design/development engineering—digital |
| 2. Engineering management | 10. Hardware engineering |
| 3. Project engineering management | 11. Software design/development |
| 4. Research and development management | 12. Computer science |
| 5. Design engineering management—analogue | 13. Science/physics/mathematics |
| 6. Design engineering management—digital | 14. Engineering (not elsewhere specified) |
| 7. Research and development engineering | 15. Marketing/sales/purchasing |
| 8. Design/development engineering—analogue | 16. Consulting |
| | 17. Education/teaching |
| | 18. Retired |
| | 19. Other _____ |

C. Principal responsibility _____ →

- | | |
|--|-----------------------|
| 1. Engineering and scientific management | 6. Education/teaching |
| 2. Management other than engineering | 7. Consulting |
| 3. Engineering design | 8. Retired |
| 4. Engineering | 9. Other _____ |
| 5. Software: science/mngmnt/engineering | |

D. Title _____ →

- | | |
|--|--------------------------------|
| 1. Chairman of the Board/President/CEO | 10. Design Engineering Manager |
| 2. Owner/Partner | 11. Design Engineer |
| 3. General Manager | 12. Hardware Engineer |
| 4. VP Operations | 13. Software Engineer |
| 5. VP Engineering/Dir. Engineering | 14. Computer Scientist |
| 6. Chief Engineer/Chief Scientist | 15. Dean/Professor/Instructor |
| 7. Engineering Management | 16. Consultant |
| 8. Scientific Management | 17. Retired |
| 9. Member of Technical Staff | 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

(continued on next page)

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.

		BETWEEN 16 AUG 2015- 28 FEB 2016 PAY	BETWEEN 1 MAR 2016- 15 AUG 2016 PAY
IEEE Aerospace and Electronic Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	AES010	25.00 <input type="checkbox"/>	12.50 <input type="checkbox"/>
IEEE Antennas and Propagation <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	AP003	15.00 <input type="checkbox"/>	7.50 <input type="checkbox"/>
IEEE Broadcast Technology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	BT002	15.00 <input type="checkbox"/>	7.50 <input type="checkbox"/>
IEEE Circuits and Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CAS004	22.00 <input type="checkbox"/>	11.00 <input type="checkbox"/>
IEEE Communications <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	COM019	30.00 <input type="checkbox"/>	15.00 <input type="checkbox"/>
IEEE Components, Packaging, & Manu. Tech. <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CPMT021	15.00 <input type="checkbox"/>	7.50 <input type="checkbox"/>
IEEE Computational Intelligence <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CIS011	29.00 <input type="checkbox"/>	14.50 <input type="checkbox"/>
IEEE Computer <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CO16	56.00 <input type="checkbox"/>	28.00 <input type="checkbox"/>
IEEE Consumer Electronics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CE008	20.00 <input type="checkbox"/>	10.00 <input type="checkbox"/>
IEEE Control Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	CS023	25.00 <input type="checkbox"/>	12.50 <input type="checkbox"/>
IEEE Dielectrics and Electrical Insulation <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	DEI032	26.00 <input type="checkbox"/>	13.00 <input type="checkbox"/>
IEEE Education <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	EO25	20.00 <input type="checkbox"/>	10.00 <input type="checkbox"/>
IEEE Electromagnetic Compatibility <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	EMC027	31.00 <input type="checkbox"/>	15.50 <input type="checkbox"/>
IEEE Electron Devices <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	ED015	18.00 <input type="checkbox"/>	9.00 <input type="checkbox"/>
IEEE Engineering in Medicine and Biology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	EMB018	40.00 <input type="checkbox"/>	20.00 <input type="checkbox"/>
IEEE Geoscience and Remote Sensing <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	GRS029	19.00 <input type="checkbox"/>	9.50 <input type="checkbox"/>
IEEE Industrial Electronics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	IE013	9.00 <input type="checkbox"/>	4.50 <input type="checkbox"/>
IEEE Industry Applications <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	IA034	20.00 <input type="checkbox"/>	10.00 <input type="checkbox"/>
IEEE Information Theory <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	IT012	30.00 <input type="checkbox"/>	15.00 <input type="checkbox"/>
IEEE Instrumentation and Measurement <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	IM009	29.00 <input type="checkbox"/>	14.50 <input type="checkbox"/>
IEEE Intelligent Transportation Systems <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	ITSS038	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Magnetics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	MAG033	26.00 <input type="checkbox"/>	13.00 <input type="checkbox"/>
IEEE Microwave Theory and Techniques <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	MTT017	17.00 <input type="checkbox"/>	8.50 <input type="checkbox"/>
IEEE Nuclear and Plasma Sciences <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	NPS005	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Oceanic Engineering <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	OE022	19.00 <input type="checkbox"/>	9.50 <input type="checkbox"/>
IEEE Photonics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PHO036	34.00 <input type="checkbox"/>	17.00 <input type="checkbox"/>
IEEE Power Electronics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PEL035	25.00 <input type="checkbox"/>	12.50 <input type="checkbox"/>
IEEE Power & Energy <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PE031	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Product Safety Engineering <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PSE043	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Professional Communication <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	PC026	31.00 <input type="checkbox"/>	15.50 <input type="checkbox"/>
IEEE Reliability <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	RL007	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Robotics and Automation <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	RA024	9.00 <input type="checkbox"/>	4.50 <input type="checkbox"/>
IEEE Signal Processing <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	SP001	22.00 <input type="checkbox"/>	11.00 <input type="checkbox"/>
IEEE Social Implications of Technology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	SIT030	33.00 <input type="checkbox"/>	16.50 <input type="checkbox"/>
IEEE Solid-State Circuits <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	SSC037	22.00 <input type="checkbox"/>	11.00 <input type="checkbox"/>
IEEE Systems, Man, & Cybernetics <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	SMC028	12.00 <input type="checkbox"/>	6.00 <input type="checkbox"/>
IEEE Technology & Engineering Management <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	TEM014	35.00 <input type="checkbox"/>	17.50 <input type="checkbox"/>
IEEE Ultrasonics, Ferroelectrics, & Frequency Control <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	UFFC020	20.00 <input type="checkbox"/>	10.00 <input type="checkbox"/>
IEEE Vehicular Technology <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>	VT006	18.00 <input type="checkbox"/>	9.00 <input type="checkbox"/>

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 2016 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 2015- 28 FEB 2016 PAY	BETWEEN 1 MAR 2016- 15 AUG 2016 PAY
United States.....	\$197.00 <input type="checkbox"/>	\$98.50 <input type="checkbox"/>
Canada (GST)*.....	\$173.35 <input type="checkbox"/>	\$86.68 <input type="checkbox"/>
Canada (NB, NF and ON HST)*.....	\$185.11 <input type="checkbox"/>	\$92.56 <input type="checkbox"/>
Canada (Nova Scotia HST)*.....	\$188.05 <input type="checkbox"/>	\$94.03 <input type="checkbox"/>
Canada (PEI HST)*.....	\$186.58 <input type="checkbox"/>	\$93.29 <input type="checkbox"/>
Canada (GST and QST Quebec).....	\$188.01 <input type="checkbox"/>	\$94.01 <input type="checkbox"/>
Africa, Europe, Middle East.....	\$160.00 <input type="checkbox"/>	\$80.00 <input type="checkbox"/>
Latin America.....	\$151.00 <input type="checkbox"/>	\$75.50 <input type="checkbox"/>
Asia, Pacific.....	\$152.00 <input type="checkbox"/>	\$76.00 <input type="checkbox"/>

*IEEE Canada Business No. 125634188

Minimum Income or Unemployed Provision

Applicants who certify that their prior year income did not exceed US\$14,700 (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\$14,700 in 2015
- I certify that I was unemployed in 2015

7 More Recommended Options

- Proceedings of the IEEE..... print \$47.00 or online \$41.00
- Proceedings of the IEEE (print/online combination).....\$57.00
- IEEE Standards Association (IEEE-SA).....\$53.00
- 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 \$ _____
- IEEE Society dues (optional) \$ _____
- IEEE-SA/WIE dues (optional) \$ _____
- Proceedings of the IEEE (optional) \$ _____
- 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.



Credit Card Number _____

MONTH _____ YEAR _____

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 _____

15-MEM-385 P 6/15

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

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.]

Digital Object Identifier

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/>

2016 IEEE SIGNAL PROCESSING SOCIETY MEMBERSHIP APPLICATION

Mail to: IEEE OPERATIONS CENTER, ATTN: Matthew Plotner, 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: _____

2016 SPS MEMBER RATES

	16 Aug-28 Feb	1 Mar-15 Aug
	Pay Full Year	Pay Half Year
Signal Processing Society Membership Fee*	\$ 22.00 <input type="checkbox"/> \$ 11.00 <input type="checkbox"/>	
<small>Fee includes: IEEE Signal Processing Magazine (electronic and digital), Inside Signal Proc. eNewsletter (electronic) and IEEE Signal Processing Society Content Gazette (electronic).</small>		
Add \$17 to enhance SPS Membership and also receive:	\$17.00 <input type="checkbox"/> \$ 8.50 <input type="checkbox"/>	
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.		
<i>Publications available only with SPS membership:</i>		
Signal Processing, IEEE Transactions on:	Print \$209.00 <input type="checkbox"/> \$104.50 <input type="checkbox"/>	
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print \$160.00 <input type="checkbox"/> \$ 80.00 <input type="checkbox"/>	
Image Processing, IEEE Transactions on:	Print \$207.00 <input type="checkbox"/> \$103.50 <input type="checkbox"/>	
Information Forensics and Security, IEEE Trans. on:	Print \$179.00 <input type="checkbox"/> \$ 89.50 <input type="checkbox"/>	
IEEE Journal of Selected Topics in Signal Processing:	Print \$176.00 <input type="checkbox"/> \$ 88.00 <input type="checkbox"/>	
Affective Computing, IEEE Transactions on:	Electronic \$ 36.00 <input type="checkbox"/> \$ 18.00 <input type="checkbox"/>	
Biomedical and Health Informatics, IEEE Journal of:	Print \$ 55.00 <input type="checkbox"/> \$ 27.50 <input type="checkbox"/>	
	Electronic \$ 40.00 <input type="checkbox"/> \$ 20.00 <input type="checkbox"/>	
	Print & Electronic \$ 65.00 <input type="checkbox"/> \$ 32.50 <input type="checkbox"/>	
IEEE Cloud Computing	Electronic and Digital \$ 39.00 <input type="checkbox"/> \$ 19.50 <input type="checkbox"/>	
IEEE Trans. on Cognitive Comm. & Networking	Electronic \$ 27.00 <input type="checkbox"/> \$ 13.50 <input type="checkbox"/>	
IEEE Trans. on Computational Imaging	Electronic \$ 28.00 <input type="checkbox"/> \$ 14.00 <input type="checkbox"/>	
IEEE Trans. on Big Data	Electronic \$ 26.00 <input type="checkbox"/> \$ 13.00 <input type="checkbox"/>	
IEEE Trans. on Molecular, Biological, & Multi-scale Communications	Electronic \$ 25.00 <input type="checkbox"/> \$ 12.50 <input type="checkbox"/>	
IEEE Internet of Things Journal	Electronic \$ 26.00 <input type="checkbox"/> \$ 13.00 <input type="checkbox"/>	
IEEE Trans. on Cloud Computing	Electronic \$ 43.00 <input type="checkbox"/> \$ 21.50 <input type="checkbox"/>	
IEEE Trans. on Computational Social Systems	Electronic \$ 30.00 <input type="checkbox"/> \$ 15.00 <input type="checkbox"/>	
IEEE Trans. on Signal & Info Proc. Over Networks	Electronic \$ 28.00 <input type="checkbox"/> \$ 14.00 <input type="checkbox"/>	
IEEE Biometrics Compendium:	Online \$ 30.00 <input type="checkbox"/> \$ 15.00 <input type="checkbox"/>	
Computing in Science & Engrg. Mag.:	Electronic and Digital \$ 39.00 <input type="checkbox"/> \$ 19.50 <input type="checkbox"/>	
	Print \$ 69.00 <input type="checkbox"/> \$ 34.50 <input type="checkbox"/>	
Medical Imaging, IEEE Transactions on:	Print \$ 74.00 <input type="checkbox"/> \$ 37.00 <input type="checkbox"/>	
	Electronic \$ 53.00 <input type="checkbox"/> \$ 26.50 <input type="checkbox"/>	
	Print & Electronic \$ 89.00 <input type="checkbox"/> \$ 44.50 <input type="checkbox"/>	
Mobile Computing, IEEE Transactions on:	ELE/Print Abstract/CD-ROM \$ 41.00 <input type="checkbox"/> \$ 20.50 <input type="checkbox"/>	
Multimedia, IEEE Transactions on:	Electronic \$ 43.00 <input type="checkbox"/> \$ 21.50 <input type="checkbox"/>	
IEEE MultiMedia Magazine:	Electronic and Digital \$ 39.00 <input type="checkbox"/> \$ 19.50 <input type="checkbox"/>	
	Print \$ 69.00 <input type="checkbox"/> \$ 34.50 <input type="checkbox"/>	
Network Science and Engrg., IEEE Trans. on:	Electronic \$ 34.00 <input type="checkbox"/> \$ 17.00 <input type="checkbox"/>	
IEEE Reviews in Biomedical Engineering:	Print \$ 25.00 <input type="checkbox"/> \$ 12.50 <input type="checkbox"/>	
	Print & Electronic \$ 40.00 <input type="checkbox"/> \$ 20.00 <input type="checkbox"/>	
IEEE Security and Privacy Magazine:	Electronic and Digital \$ 39.00 <input type="checkbox"/> \$ 19.50 <input type="checkbox"/>	
	Print \$ 69.00 <input type="checkbox"/> \$ 34.50 <input type="checkbox"/>	
IEEE Sensors Journal:	Electronic \$ 40.00 <input type="checkbox"/> \$ 20.00 <input type="checkbox"/>	
Smart Grid, IEEE Transactions on:	Print \$100.00 <input type="checkbox"/> \$ 50.00 <input type="checkbox"/>	
	Electronic \$ 40.00 <input type="checkbox"/> \$ 20.00 <input type="checkbox"/>	
	Print & Electronic \$120.00 <input type="checkbox"/> \$ 60.00 <input type="checkbox"/>	
Wireless Communications, IEEE Transactions on:	Print \$127.00 <input type="checkbox"/> \$ 63.50 <input type="checkbox"/>	
	Electronic \$ 49.00 <input type="checkbox"/> \$ 24.50 <input type="checkbox"/>	
	Print & Electronic \$127.00 <input type="checkbox"/> \$ 63.50 <input type="checkbox"/>	
IEEE Wireless Communications Letters:	Electronic \$ 19.00 <input type="checkbox"/> \$ 9.50 <input type="checkbox"/>	

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. 2016 IEEE MEMBER RATES

IEEE DUES Residence	16 Aug 14-28 Feb 15	1 Mar -15 Aug 15
	Pay Full Year	Pay Half Year**
United States	\$197.00 <input type="checkbox"/>	\$98.50 <input type="checkbox"/>
Canada (incl. GST)	\$173.35 <input type="checkbox"/>	\$86.68 <input type="checkbox"/>
Canada (incl. HST for PEI)	\$186.58 <input type="checkbox"/>	\$93.29 <input type="checkbox"/>
Canada (incl. HST for Nova Scotia)	\$188.05 <input type="checkbox"/>	\$94.03 <input type="checkbox"/>
Canada (incl. HST for NB, NF and ON)	\$185.11 <input type="checkbox"/>	\$92.56 <input type="checkbox"/>
Canada (incl. GST and GST Quebec)	\$188.01 <input type="checkbox"/>	\$94.01 <input type="checkbox"/>
Africa, Europe, Middle East	\$160.00 <input type="checkbox"/>	\$80.00 <input type="checkbox"/>
Latin America	\$151.00 <input type="checkbox"/>	\$75.50 <input type="checkbox"/>
Asia, Pacific	\$152.00 <input type="checkbox"/>	\$76.00 <input type="checkbox"/>

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 12%, 13% 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 EU926000081)

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.



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): _____

