

IEEE **Signal Processing** MAGAZINE

Volume 34 | Number 2 | March 2017

SMART VEHICLE TECHNOLOGIES: PART 2

Signal Processing on the Move

Perfecting Protection
for Interactive Multimedia

Computational Neuromodulation
for Deep Brain Stimulation

New Fellows and
SPS Award Winners



General Chairs

Warren Gross
McGill Univ.

Kostas Plataniotis
Univ. of Toronto

Technical Program Chairs

Z. Jane Wang
Univ. of British Columbia

Costas Kotropoulos
Univ. of Thessaloniki

Qionghai Dai
Tsinghua Univ., China

Finance Chair

Ioannis Psaromiligkos
McGill Univ.

Local Arrangement Chair

Fabrice Labeau
McGill Univ.

Publications Chair

Xiangyang Ji
Tsinghua Univ., China

Publicity Chairs

Purang Abolmaesumi
Univ. of British Columbia

Student Program Chair

Ivan Bajic
Simon Fraser University

Industrial Program Chairs

André Morin
Optelis

Panos Nasiopoulos
Univ. of British Columbia

International Liaison Chair

Chunyan Miao
Nanyang Tech. Univ.

Conference Board Liaison

Dapeng Wu
Univ. of Florida

Advisory Committee Chair

Rabab Ward
Univ. of British Columbia

IEEE GlobalSIP



5th IEEE Global Conference on Signal and Information Processing

November 14 - 16, 2017, Montreal, Canada



www.ieeeglobalsip.org/

Call for Papers

The 5th IEEE Global Conference on Signal and Information Processing (GlobalSIP) will be held in Montreal, Quebec, Canada on November 14-16, 2017. GlobalSIP, as a new flagship IEEE Signal Processing Society conference, focuses on signal and information processing with an emphasis on up-and-coming signal processing themes. The conference features world-class plenary speeches, distinguished Symposium talks, tutorials, exhibits, oral and poster sessions, and panels. GlobalSIP is comprised of co-located General Symposium and symposia selected based on responses to the call-for-symposia proposals. Featured symposia include:

- General symposium
- Sparse SP and deep learning
- Graph signal processing
- Distributed optimization and resource management over networks
- Control & information theoretic approaches to privacy and security
- SP for interference cancellation and full-duplex communication systems
- SP for Accelerating Deep Learning
- SP for Smart Cities & IoTs
- SP & ML in large medical datasets
- Big Data Analytics for IoT Healthcare
- Advanced Bio-SP for Rehabilitation and Assistive Systems
- Deep Learning for Intelligent Multimedia
- Knowledge-based Multimedia Computing
- Stochastic & approximate computing for signal processing and machine learning
- Signal and information processing (SIP) for smart grid infrastructure
- SIP for Healthcare Engineering
- SIP for Finance and Business

Prospective authors are invited to submit full-length papers (up to 4 pages for technical content, an optional 5th page containing only references) and extended abstracts (up to 2 pages, for paper-less industry presentations and Ongoing Work presentations). Manuscripts should be original and written in accordance with the standard IEEE 2-column paper template. Accepted full-length papers will be indexed on IEEE Xplore. Accepted abstracts will not be indexed in IEEE Xplore, however the abstracts and/or the presentations will be included in the IEEE SPS SigPort. Accepted papers and abstracts will be scheduled in lecture and poster sessions.

Conference Highlights

- Plenary Talks within the general symposium and Distinguished Symposium Presentations/Talks within the 16 thematic symposia surveying emerging topics in SIP
- Panel discussions on funding opportunities, trends and targeted topics
- Enhanced industry program: paperless industry presentations, panels, demos and exhibitions
- Exciting student program: Ongoing Work tracks, Student-Industry Luncheon, 3MT competition
- Great venue with vibrant cultural, educational, and scientific identity, combining the modern buzz of a North American city and a specific European charm

- | | | |
|-------------------------|-----------------------|--|
| Important Dates: | • May 15, 2017 | : Paper submission due |
| | • June 30, 2017 | : Final acceptance notifications sent to all authors |
| | • July 22, 2017 | : Camera-ready papers due |



Contents

Volume 34 | Number 2 | March 2017

SPECIAL SECTION

SIGNAL PROCESSING FOR SMART VEHICLE TECHNOLOGIES: PART 2

18 FROM THE GUEST EDITORS

John H.L. Hansen, Kazuya Takeda, Sanjeev M. Naik, Mohan M. Trivedi, Gerhard U. Schmidt, Yingying (Jennifer) Chen, and Wade Trappe

22 AUTOMOTIVE RADARS

Sujeet Patole, Murat Torlak, Dan Wang, and Murtaza Ali

36 ADVANCES IN AUTOMOTIVE RADAR

Florian Engels, Philipp Heidenreich, Abdelhak M. Zoubir, Friedrich K. Jondral, and Markus Wintermantel

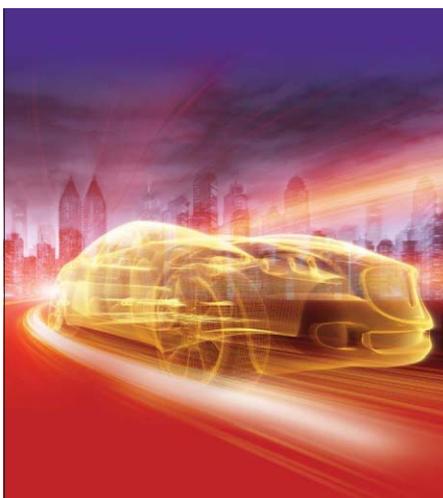
47 SIGNAL PROCESSING CHALLENGES IN CELLULAR-ASSISTED VEHICULAR COMMUNICATIONS

Stefan Schwarz, Tal Filosof, and Markus Rupp

IMAGE LICENSED BY GRAPHIC STOCK



PG. 124



ON THE COVER

Part 2 of the special section on signal processing for smart vehicle technologies provides an additional six articles, complementing those in Part 1, which was published in the November 2016 issue of *IEEE Signal Processing Magazine*.

COVER IMAGE: ©ISTOCKPHOTO.COM/NADIA

60 THE FUTURE OF AUTOMOTIVE LOCALIZATION ALGORITHMS

Rickard Karlsson and Fredrik Gustafsson

70 DSP APPLICATIONS IN ENGINE CONTROL AND ONBOARD DIAGNOSTICS

Zhijian James Wu and Sanjeev M. Naik

82 INTELLIGENT INTERACTIVE DISPLAYS IN VEHICLES WITH INTENT PREDICTION

Bashar I. Ahmad, James K. Murphy, Simon J. Godsill, Patrick M. Langdon, and Robert W. Hardy

FEATURE

95 PERFECTING PROTECTION FOR INTERACTIVE MULTIMEDIA

Ahmed Badr, Ashish Khisti, Wai-Tian Tan, and John Apostolopoulos

COLUMNS

7 Society News

SPS Fellows and Award Winners Recognized

12 Reader's Choice

Top Downloads in *IEEE Xplore*

14 Special Reports

Stepping Up Security with Signal Processing
John Edwards

114 Life Sciences

Computational Neuromodulation: Future Challenges for Deep Brain Stimulation
Konstantinos P. Michmizos, Blerita Lindqvist, Stephen Wong, Eric L. Hargreaves, Konstantinos Psychas, Georgios D. Mitsis, Shabbar F. Danish, and Konstantina S. Nikita

124 In the Spotlight

No Need for Speed: More Signal Processing Innovation Is Required Before Adopting Automated Vehicles
Wade Trappe



PG. 14

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\$241.00 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 © 2017 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.2016.2642406

DEPARTMENTS

3 From the Editor

A Conversation About Signal Processing at Elementary School
Min Wu

5 President's Message

Introducing the IEEE Signal Processing Society Executive Committee
Rabab Ward

123 Dates Ahead



The IEEE International Symposium on Biomedical Imaging will be held 18–21 April in Melbourne, Australia.



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

IEEE Signal Processing Magazine

EDITOR-IN-CHIEF

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

AREA EDITORS

Feature Articles

Shuguan 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

Mrityunjay Chakraborty—Indian Institute of
Technology, Kharagpur, India

George Christos—Qualcomm, Inc.,
U.S.A.

Alfonso Farina—Selex ES, Italy

Mounir Ghogho—University of Leeds,
U. K.

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

C.-C. Jay Kuo—University of Southern California,
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.

Vahid Tarokh—Harvard University, U.S.A.

Wade Trappe—Rutgers University, U.S.A.

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

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

Changshui Zhang—Tsinghua University, China

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

Nikos D. Sidiropoulos—Vice President,
Membership

Thrasyloulos (Thrasos) N. Pappas—
Vice President, Publications

Walter Kellerman—Vice President,
Technical Directions

IEEE SIGNAL PROCESSING SOCIETY STAFF

Rebecca Wollman—Publications Administrator

IEEE PERIODICALS MAGAZINES DEPARTMENT

Jessica Barragué, *Managing Editor*

Geraldine Krolin-Taylor,
Senior Managing Editor

Mark David, *Director, Business Development -
Media & Advertising*

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.2017.2650178

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

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

A Conversation About Signal Processing at Elementary School

When you receive this issue of *IEEE Signal Processing Magazine (SPM)*, a number of readers will be heading to New Orleans, Louisiana, for the 42nd International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017). New Orleans is well known for its cross-cultural heritage embodied by its distinctive architecture, cuisine, and music. I look forward to this annual gathering and flagship conference of members of the IEEE Signal Processing Society (SPS), and I hope to see our magazine team and many readers there. A number of SPS boards and committees will meet at ICASSP. Among them, several are related to education and outreach for which the SPS has been making significant efforts to improve the understanding and appreciation of our field by outside communities.

Speaking of education, the public school system where I live has a tradition of encouraging parents' participation in their children's learning. One of the weekly activities that my son's elementary school invites parents to sign up for is called "Mystery Reader." This is a fun way for parents to spend some time reading a book or other appropriate material of their choice to the class. Adding to the mystery is the surprise factor: a parent does not share with his/her child that he/she is the mystery reader and instead provides a clue to the teacher who will then share the clue with the class and ask them to guess whose parent is coming to read.

Digital Object Identifier 10.1109/MSP.2017.2650179
Date of publication: 3 March 2017

When I did this last year, I picked a book from the "Who Was/Is?" series of biographies published by Random House tailored to elementary school readers. I read to students excerpts of *Who Was Dr. Seuss?* (by Janet Pascal and illustrated by Nancy Harrison). As some of our readers may know, Dr. Seuss is the pen name of Theodor Seuss Geisel, an American writer, cartoonist, and artist who authored some of the most popular children's books, many of which have been used in American elementary classrooms today. One of his most beloved books, *Green Eggs and Ham*, consists of just 50 different words from a simple vocabulary for beginning readers. With bouncy rhymes, an imaginary storyline, and fun visual characters, Dr. Seuss revolutionized children's reading material from the quite limited and boring collection of the past.

This time, I was inspired after seeing my son's class "publish" a book nearly every year, whereby students developed every page with a drawing and a few lines of writing. I thought it would be fun to bring something that I wrote during my studies to share with the class, such as *Multimedia Data Hiding* (Springer, 2003, with Prof. Bede Liu), a book extended from my doctoral thesis work; I also picked two issues of *SPM*, to serve as more recent examples.

As I entered the classroom, over two dozen third graders sat around a rocking chair with curiosity and eagerness on their faces. I started by showing them the book's cover, explaining in as simple language as I could about this research of putting invisible data in pictures. I then

took a highly sparse "sampling" of a few pages in the book. I showed them a cartoonish block diagram outlining the data embedding process, a pictorial example to visualize the before and after of data embedding and extraction, and a page containing an illustration of groupings of colored balls drawn randomly from a bag and many equations modeling this game. These examples showed the students that grown-up researchers try various ways to explain their work to readers, and their skills are built on top of what they learned in elementary school.

Moving to the second half of my reading, I took out the two issues of *SPM*. When I explained the name of the magazine, which is also the field of our study, I asked the students whether they had heard of signal processing and any examples they knew. A girl sitting in the front raised her hand high: "It is about radar and the signals it has ..." Thanks to a number of R&D labs and companies in the Greater Washington, D.C. area, some of our young people have picked up an amazing amount of technical terms. "Also the pictures! And music! And cell phones!" said several students, eager to contribute. Their minds are like a sponge—once stimulated, they are so absorptive that they had already related to what we had discussed a short while ago on the digital revolution and the gadgets they have.

I showed them the cover of *SPM*'s September 2016 special issue on researching new generations of cameras and displays. I flipped to the the guest editorial [1], a page containing a visual summary of the articles included in the

special issue. There, they saw a miniature camera that is smaller than a fingernail, a light-field camera that captures different focus and depth of a scene, a dissection of a virtual reality (VR) headset, and an analysis of the invisible signatures in paintings, just to name a few. The class was excited about these: “Oh, yes, I tried new goggles, and it’s so cool!” said a boy about a VR headset he tested during the holiday season. “I have to have that camera,” said a girl to me afterward.

I then turned to the first page of the magazine and told them that I had an essay assignment for each magazine issue, just like the writing tasks they regularly have, and that my editorials drew inspirations from interacting with other people, including them—my son’s elementary school class. To calm the puzzled expressions on many of the students’ faces, I read to them the first part of my editorial in the March 2015 issue of the magazine [2]. That particular editorial started by me sharing holiday traditions and cultural heritages that were done so

well in my son’s elementary school. Then the editorial discussed the sharing of our technical field with the general public. There I was inspired by the computing community’s one-hour coding effort, including the enthusiastic reaction from the outreach at my son’s school.

At the end of this reading session, I thanked the class for letting me share my technical area with them. To be honest, it was not an easy job in the traditional format of reading and conversation, without using a projector, slides, or board as in a professional presentation and teaching that many of us have gotten so used to. Yet, these young students participated with full curiosity and enthusiasm. A girl approached me right after and said, “You definitely should talk to our technology teacher,” (who teaches them the basics about computing, including one-hour coding as early as kindergarten). She paused and added, “How about writing a children’s book, a picture book about your signals, maybe about how it got the processing by that cool camera?” Well,

while we might have thought that young minds would prefer video to print material, we as a community have just received a challenge from an eight-year-old: a signal processing book for children!

Many of us have seen, with awe and, maybe, also a hint of jealousy, the high profile of the computer coding initiatives. If Dr. Seuss could win a bet with his publisher over the seemingly impossible task of using just 50 simple words to write a fun story, perhaps there is some truth to this surprise “assignment” that is worth our pondering!

References

- [1] A. Agrawal, R. Baraniuk, P. Favaro, and A. Veeraraghavan, “Signal processing for computational photography and displays [From the Guest Editors],” *IEEE Signal Process. Mag.* vol. 33, no. 5, pp. 12–15, Sept. 2016.
- [2] M. Wu, “Sharing signal processing with the world [From the Editor],” *IEEE Signal Process. Mag.* vol. 32, no. 2, p. 4, Mar. 2015.

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 2017!

- **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.
- **How?** Login on www.sigport.org using IEEE web account credentials. Go to “submit your work” on the top menu and use promotion code you14200 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 through www.sigport.org/contact.



Digital Object Identifier 10.1109/MSP.2017.2673078

PRESIDENT'S MESSAGE

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

Introducing the IEEE Signal Processing Society Executive Committee

Much of the everyday business of the IEEE Signal Processing Society (SPS) runs like a well-oiled machine, with SPS staff in Piscataway, New Jersey, fielding requests and managing emergencies from volunteers and members, editors and authors, students and fellows. The SPS executive staff is integral to running the Society's business, whether it's managing the peer-review process for our publications, updating the SPS website, or launching marketing campaigns to attract new members. However, they are also essential to what makes SPS a leader within the IEEE and signal processing professionals—executing the vision of the SPS Board of Governors, as directed and implemented by the SPS Executive Committee.

The SPS Executive Committee is made up of six individuals, spanning various technical interests, geographic regions, and Society activities: Conferences, Membership, Publications, and Technical Directions. Each of these areas has a vice president, all falling under the leadership of the SPS president and working alongside the SPS president-elect to look forward and establish the strategic vision, direction, and impact of the Society for years to come. In addition, each vice president has several subcommittees that enable Society programs and capture the collaborative spirit of SPS volunteerism. While getting all of

these—sometimes conflicting—ideas, products, and people to work together can be a challenge, the level of cooperation and synergy required is both undeniable and rewarding.

Conferences are the primary driver of SPS business, and Carlo Regazzoni of the University of Genoa, Italy, leads the SPS Conference Board in his role of vice president—Conferences. With three flagship conferences, dozens of financially and technically sponsored and cosponsored workshops, exploring opportunities for growth, watching for new trends in technical meetings, and ensuring that the SPS stays at the forefront with exciting events, SPS conferences are quite a task. A single conference takes many years, and many people, to plan successfully. Managing dozens of budgets, venues, organizing committees, and symposia is a demanding task, requiring a level of meticulousness, patience, and foresight among the dozens of involved parties. However, if you've ever attended an SPS conference, you have experienced the reward firsthand: SPS conferences are world-class events, attended by thousands annually. Not only do they provide an opportunity to learn and share the highest caliber of research and knowledge, but they are vibrant exchanges of information, a chance to network with colleagues, a chance to collaborate, build relationships, and grow our worldwide signal processing community.

The position of vice president—Membership is the newest addition to the SPS

Executive Committee and has become increasingly significant since it was established in 2013. In January, we welcomed Nikos Sidiropoulos of the University of Minnesota into the position. Many standing committees sit under the SPS Membership Board, and, naturally, SPS membership activities are the heart of the SPS member experience. This board, like SPS membership, is broad and diverse, serving a growing and evolving member base with changing needs and values. From 183 Chapters worldwide to student activities, to educational programs and seasonal schools, the multifaceted nature of the Membership Board and its services presents a unique set of challenges to remain relevant and valuable in a rapidly changing world of association membership and the professionals they support. Their impact is apparent in every graduate exposed to the breadth of available signal processing careers at a student career luncheon; in every student attending an SPS conference on an SPS travel grant; and in every connection made and community built at a local Chapter event, Young Professionals workshop, or Women in Signal Processing luncheon.

SPS publications are the cornerstone of SPS educational materials, largely responsible for the Society's longstanding esteem and authority among signal processing professionals. Our vice president—Publications, Thrasyvoulos Pappas of Northwestern University, Evanston, Illinois, currently leads the SPS Publications

Digital Object Identifier 10.1109/MSP.2017.2650180
Date of publication: 3 March 2017

Board, made up of nine editors-in-chief of SPS publications alongside the SPS vice presidents of Conferences and Technical Directions, assuring that SPS's periodicals (including those that are jointly published with other Societies) are meeting the high standard of quality, timeliness, and technical relevance that SPS members and customers have come to expect out of our research. Looking ahead, the SPS vice-president–Publications plans how to keep the delivery of our publications modern and timely through our ever-expanding technological landscape and also ensures that our content, and the subject areas presented within it, remain relevant to the greater signal processing community. Whether it be the more traditional papers found in our journals, conference proceedings in *IEEE Xplore*, or emerging news in *IEEE Signal Processing Magazine*, SPS publications and the editorial boards that manage them honor the Society and our field with high-quality, trusted materials that have continued to cement the SPS as the leader among the many industries, applications, and technologies we enable.

And finally, the SPS Technical Directions Board ties it all together. After

all, engineering, technology, and the advancement of each are at the heart of the IEEE's mission. Chaired by Walter Kellermann of the University Erlangen-Nuremberg in Germany, the Technical Directions Board comprises the chairs of SPS's 15 technical committees (TCs) and special interest groups (SIGs), as well as various other board and committee chairs across Conferences, Publications, and Membership, to ensure that all of the SPS's products, services, and activities are in sync; aligned with our technical vision; and delivering the level of quality that our members and the signal processing community have come to expect. The vice president–Technical Directions is not only concerned with future directions of the state of the art for existing technologies but also with new emerging technologies that relate to our field. The Technical Directions Board, along with our TCs and SIGs, have a resounding voice in all technical content, including conference symposia, publications, awards, and education programs. Their input is imperative to all Society activities and allowing us to remain on the cutting edge within the IEEE and the overall engineering community.

As a unit, the SPS Executive Committee forms a forward-thinking team to manage the Society's products, programs, and benefits to set the course for the Society's long-term strategic vision. Alongside the Board of Governors, the SPS Executive Committee identifies and aligns objectives to prepare for not only the next generation of SPS members but future generations of signal processing engineers who expect our products and services to reflect the pace innovation of the field itself. I am honored to work with SPS President-Elect Ali H. Sayed to prioritize and guide these objectives and initiatives into achievements that will serve the Society and its members for years to come. Whether the future is focused on public visibility, continuing education, industry engagement, or professional development, we are all working together to—like signal processing as a field—lay the groundwork for a better tomorrow.




Nominate a deserving colleague

- **SPS Awards Nomination:** Generally in June to August each year. Please see the society's announcement and contact technical committees for their respective endorsement schedule. <http://signalprocessingsociety.org/get-involved/awards-submit-award-nomination>
- **IEEE Medals:** Nomination is generally due in June each year. Visit the websites for details. <https://www.ieee.org/about/awards/medals/index.html>
- **IEEE Technical Field Awards (TFA):** Nomination is generally due in January each year. Visit the TFA websites for details. https://www.ieee.org/about/awards/tfas/index_tfas.html
- **IEEE Fellow:** Nomination is generally due on March 1 each year. Please visit the fellow program website for details. <http://www.ieee.org/fellowprogram>

Digital Object Identifier 10.1109/MSP.2017.2673080

SPS Fellows and Award Winners Recognized

In this column of *IEEE Signal Processing Magazine*, 47 IEEE Signal Processing Society (SPS) members are recognized as IEEE Fellows and award recipients are announced.

47 SPS members elevated to Fellow

Each year, the IEEE Board of Directors confers the grade of Fellow on up to one-tenth of 1% of the Members. To qualify for consideration, an individual must have been a Member, normally for five years or more, and a Senior Member at the time of nomination to Fellow. The grade of Fellow recognizes unusual distinction in IEEE's designated fields.

The SPS congratulates the following 47 SPS members who were recognized with the grade of Fellow as of 1 January 2017.

Raviraj Adve, Toronto, Canada, for development of signal processing techniques for airborne radar.

Sos Agaian, San Antonio, Texas, for contributions to biologically inspired visual data processing systems.

R. Michael Buehrer, Blacksburg, Virginia, for contributions to wideband signal processing in communications and geolocation.

Lap-Pui Chau, Singapore, for contributions to fast computation algorithms for visual signal processing.

Douglas Cochran, Tempe, Arizona, for contributions to multichannel coherence in radar, sonar, and spectrum sensing.

Huaiyu Dai, Raleigh, North Carolina, for contributions to multiple-input, multiple-output communications and wireless security.

Ricardo De Queiroz, Brasilia, Brazil, for contributions to image and video signal enhancement and compression.

Jose Dias, Lisboa, Portugal, for contributions to imaging inverse problems in remote sensing.

Pier Luigi Dragotti, London, United Kingdom, for contributions to sparse signal representation and sampling theory.

Pablo Estevez, Santiago, Chile, for contributions to feature selection and visualization of large data sets.

James Fowler, Mississippi State, Mississippi, for contributions to lossy source coding and dimensionality reduction of multidimensional data.

Michael Gastpar, Lausanne, Switzerland, for contributions to network information theory.

Uwe Hanebeck, Waldbronn, Germany, for contributions to nonlinear estimation and control.

Julia Hirschberg, New York, for contributions to text-to-speech synthesis and spoken language understanding.

Yo-Sung Ho, Gwangju, South Korea, for contribution to video coding and three-dimensional image processing.

Tatsuya Kawahara, Kyoto, Japan, for contributions to speech recognition and understanding.

Takao Kobayashi, Yokohama, Japan, for contributions to expressive speech synthesis based on a statistical parametric approach.

Witold Krzymien, Edmonton, Canada, for contributions to radio resource management for cellular systems and networks.

Li Li, Beijing, China, for contributions to intelligent transportation systems and vehicles.

Xin Li, Morgantown, West Virginia, for contributions to image coding, restoration, and interpolation.

Yuanqing Li, Guangzhou, China, for contributions to brain signal analysis and brain computer interfaces.

Qilian Liang, Arlington, Texas, for contributions to interval type-2 fuzzy logic systems.

Teng Joon Lim, Singapore, for contributions to statistical signal processing in wireless communications.

Wing-kin Ma, Hong Kong, for contributions to optimization in signal processing and communications.

Enrico Magli, Torino, Italy, for contributions to compression and communication of remotely sensed imagery.

Aleksandra Mojsilovic, Yorktown Heights, New York, for contributions to signal processing for image analysis, data mining, and business analytics.

Srikantan Nagarajan, San Francisco, California, for contributions to neural engineering and biomagnetic brain imaging.

Panos Nasiopoulos, Vancouver, Canada, for leadership in DVD authoring and digital multimedia technologies.

Brett Ninness, Newcastle, Australia, for contributions to computational methods in system identification.

Phillip Pace, San Juan Bautista, California, for leadership in radar signal

processing, receiver design, and direction finding architectures.

Bhiksha Raj, Pittsburgh, Pennsylvania, for contributions to speech recognition.

Bhuvana Ramabhadran, Mount Kisco, New York, for contributions to speech recognition and language processing.

Gael Richard, Paris, France, for contributions to analysis, indexing, and decomposition of audio and music signals.

Seb Savory, Ely, United Kingdom, for contributions to digital coherent transceivers for optical fiber communication.

Behzad Shahraray, Middletown, New Jersey, for leadership in content-based processing and retrieval of multimedia information.

Steven Smith, Lexington, Massachusetts, for contributions to statistical signal processing and applications to radar and sonar.

Leif Sörnmo, Lund, Sweden, for contributions to biomedical signal processing in cardiac applications.

Yannis Stylianou, Cambridge, United Kingdom, for contributions to speech analysis and communication.

Guaning Su, Singapore, for leadership in defense technology and management of educational institutions.

Sabine Susstrunk, Lausanne, Switzerland, for contributions to computational imaging, color image processing, and color computer vision.

Nuno Vasconcelos, La Jolla, California, for contributions to computer vision, image processing, and multimedia.

Xianbin Wang, London, Canada, for contributions to orthogonal frequency-division multiplexing systems and distributed transmission technologies.

Zhen Wang, Vancouver, Canada, for contributions to statistical signal processing for multimedia security and brain data analytics.

Ying Wu, Evanston, Illinois, for contributions to motion analysis and pattern discovery in computer vision.

Shuicheng Yan, Singapore, for contributions to subspace learning and visual classification.

Rui Zhang, Singapore, for contributions to cognitive radio and energy harvesting communications.

Ce Zhu, Chengdu, China, for contributions to video coding and communications.

2016 IEEE SPS Awards presented in New Orleans, Louisiana

The IEEE SPS congratulates the following SPS members who will receive the Society's prestigious awards during ICASSP 2017 in New Orleans, Louisiana.



P.P. Vaidyanathan

The Society Award honors outstanding technical contributions in a field within the scope of the IEEE SPS and outstanding leadership within that field. The Society Award comprises a plaque, a certificate, and a monetary award of US\$2,500. It is the highest-level award bestowed by the IEEE SPS. This year's recipient is P.P. Vaidyanathan, "for pioneering contributions to signal processing theory and education."



Jelena Kovačević



Bhaskar D. Rao

The Technical Achievement Award is presented this year to Jelena Kovačević "for contributions to the theory and practice of signal representations" and to Bhaskar D. Rao "for fundamental contributions to array processing and sparsity-based signal processing." The Technical Achievement Award honors a person who, over a period of years, has made outstanding technical contributions to theory and/or practice in technical areas within the scope of the Society, as demonstrated by publications, patents, or recognized impact on this field. The prize for the award is US\$1,500, a plaque, and a certificate.



K.J. Ray Liu

The Meritorious Service Award is presented this year to K.J. Ray Liu "for exemplary service to and leadership in the IEEE Signal Processing Society." The award comprises a

plaque and a certificate; judging is based on dedication, effort, and contributions to the Society.

The Sustained Impact Paper Award shall honor the author(s) of a journal article of broad interest that has had sustained impact over many years on a subject related to the Society's technical scope. The prize shall consist of US\$500 per author (up to a maximum of US\$1,500 per award) and a certificate. In the event that there are more than three authors, the maximum prize shall be divided equally among all authors, and each shall receive a certificate. To be eligible for consideration, an article must have appeared in one of the IEEE SPS transactions or *IEEE Journal of Selected Topics in Signal Processing*, in an issue predating the spring awards board meeting by at least ten years (typically held in conjunction with ICASSP). This year, the Sustained Impact Paper Award recipients are Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, April 2004.

The Overview Paper Award honors the author(s) of a journal article of broad interest that has had substantial impact over several years on a subject related to the Society's technical scope. A paper considered for the award should present an overview of a method or theory with technical depth and application perspective. It should have a multiyear record of impact and also be relevant to current researchers and/or practitioners. The prize shall consist of US\$500 per author (up to a maximum of US\$1,500 per award) and a certificate. This year, the Overview Paper Award recipients are Marco F. Duarte and Yonina C. Eldar, for "Structured Compressed Sensing: From Theory to Applications," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, September 2011.

The IEEE Signal Processing Magazine Best Paper Award honors the author(s) of an article of exceptional merit and broad interest on a subject related to the Society's technical scope and appearing in the Society's magazine.

The prize comprises US\$500 per author (up to a maximum of US\$1,500 per award) and a certificate. In the event that there are more than three authors, the maximum prize shall be divided equally among all authors, and each shall receive a certificate. This year, the IEEE Signal Processing Magazine Best Paper Award recipients are David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst for their article “The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains,” published in *IEEE Signal Processing Magazine*, vol. 30, no. 3, May 2013.

The IEEE Signal Processing Magazine Best Column Award honors the author(s) of a column of exceptional merit and broad interest on a subject related to the Society’s technical scope and appearing in the Society’s magazine. The prize shall consist of US\$500 per author (up to a maximum of US\$1,500 per award) and a certificate. In the event that there are more than three authors, the maximum prize shall be divided equally among all authors, and each shall receive a certificate. This year, the IEEE Signal Processing Magazine Best Column Award recipient is Ronald W. Schafer for his article “What Is a Savitzky-Golay Filter?” published in *IEEE Signal Processing Magazine*, vol. 28, no. 4, July 2011.

Five Best Paper Awards were given, honoring the author(s) of a paper of exceptional merit dealing with a subject related to the Society’s technical scope and appearing in one of the Society’s transactions, irrespective of the author’s age. The prize is US\$500 per author (up to a maximum of US\$1,500 per award) and a certificate. Eligibility is based on a five-year window preceding the year of election, and judging is based on general quality, originality, subject matter, and timeliness. Up to six Best Paper Awards may be presented each year. This year, the awardees are

- Jun Yu, Yong Rui, and Dacheng Tao, for “Click Prediction for Web Image Reranking Using Multimodal Sparse Coding,” *IEEE Transactions on Image Processing*, vol. 23, no. 5, May 2014.

- Chin Keong Ho, and Rui Zhang, for “Optimal Energy Allocation for Wireless Communications with Energy Harvesting Constraints,” *IEEE Transactions on Signal Processing*, vol. 60, no. 9, September 2012.

- Ossama Abdel-Hamid, Abdel-rahman Mohamed, Hui Jiang, Li Deng, Gerald Penn, and Dong Yu, “Convolutional Neural Networks for Speech Recognition,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 10, October 2014.

- Cees H. Taal, Richard C. Hendriks, Richard Heusdens, and Jesper Jensen, “An Algorithm for Intelligibility Prediction of Time–Frequency Weighted Noisy Speech,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 7, September 2011.

- Jason T. Parker, Philip Schniter, and Volkan Cevher (two-part paper), “Bilinear Generalized Approximate Message Passing—Part I: Derivation”; “Part II: Applications,” *IEEE Transactions on Signal Processing*, vol. 62, no. 22, November 2014.

The Young Author Best Paper Award honors the author(s) of an especially meritorious paper dealing with a subject related to the Society’s technical scope and appearing in one of the Society’s transactions and who, upon date of submission of the paper, is younger than 30 years of age. The prize is US\$500 per author (up to a maximum of US\$1,500 per award) and a certificate. Eligibility is based on a three-year window preceding the year of election, and judging is based on general quality, originality, subject matter, and timeliness. Four Young Author Best Paper Awards are being presented this year:

- Ahmed Alkhateeb and Omar El Ayach, for the paper coauthored with Geert Leus and Robert W. Heath, Jr., “Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, October 2014.

- Saiprasad Ravishankar, for the paper coauthored with Yoram Bresler, “Learning Sparsifying Transforms,”

IEEE Transactions on Signal Processing, vol. 61, no. 5, March 2013.

- Jack Ho, for the paper coauthored with Wee Peng Tay, Tony Q.S. Quek, and Edwin K.P. Chong, “Robust Decentralized Detection and Social Learning in Tandem Networks,” *IEEE Transactions on Signal Processing*, vol. 63, no. 19, October 2015.

- Yuanming Shi and Brendan O’Donoghue, for the paper coauthored with Jun Zhang and Khaled B. Letaief, “Large-Scale Convex Optimization for Dense Wireless Cooperative Networks,” *IEEE Transactions on Signal Processing*, vol. 63, no. 18, September 2015.

The IEEE Signal Processing Letters Best Paper Award honors the author(s) of a letter article of exceptional merit and broad interest on a subject related to the Society’s technical scope and appearing in *IEEE Signal Processing Letters*. The prize shall consist of US\$500 per author (up to a maximum of US\$1,500 per award) and a certificate. To be eligible for consideration, an article must have appeared in *IEEE Signal Processing Letters* in an issue predating the spring awards board meeting by five years (typically held in conjunction with ICASSP). Judging shall be on the basis of the technical novelty, the research significance of the work, and quality and effectiveness in presenting subjects in an area of high impact to the Society’s members. The recipients of the IEEE Signal Processing Letters Best Paper Award are Ye Yang, Qiang Li, Wing-Kin Ma, Jianhua Ge, and P.C. Ching, for “Cooperative Secure Beamforming for AF Relay Networks with Multiple Eavesdroppers,” *IEEE Signal Processing Letters*, vol. 20, no. 1, January 2013.

2016 Chapter of the Year Award

The IEEE SPS Gujarat Chapter has been selected as the sixth recipient of the 2016 Chapter of the Year Award, which will be presented at the ICASSP 2017 Awards Ceremony in New Orleans, Louisiana. The award is presented annually to a Chapter that has provided its membership with the highest quality of programs, activities, and services. The SPS

Gujarat Chapter will receive a certificate and a check in the amount of US\$1,000 to support local Chapter activities. The Chapter will publish an article in a future issue of *IEEE Inside Signal Processing e-Newsletter*.

SPS members receive 2017 IEEE awards

The IEEE has announced the recipients of the 2017 IEEE medals. IEEE medals are the highest honor of awards presented by the IEEE. The medals will be presented at the 2017 IEEE Honors Ceremony. Three SPS members were awarded with IEEE medals for 2017.



Martin Vetterli

The IEEE Jack S. Kilby Signal Processing Medal awarded for outstanding achievements in signal processing will be presented to Martin Vetterli (EPFL, Lausanne, Switzerland) “for fundamental contributions to advanced sampling, signal representations, and multirate and multi-resolution signal processing.”



H. Vincent Poor

The IEEE Alexander Graham Bell Medal awarded for exceptional contributions to communications and networking sciences and engineering, will be presented to H. Vincent Poor (Princeton University, New Jersey) “for fundamental contributions to signal processing and its application to digital communications.”



John Baras

The IEEE Simon Ramo Medal awarded for exceptional achievement in systems engineering and systems science, will be presented to John Baras (University of Maryland, College Park) “for exceptional contributions to the conception and commercialization of Internet-over-satellite systems, and for leadership in model-based engineering, systems science, and engineering research.”

In addition, the IEEE has announced the recipients of Technical Field Awards for contributions or leadership in specific fields of interest of the IEEE. The following SPS members have received IEEE Technical Field Awards this year.



Mark Yoffe Liberman

The IEEE James L. Flanagan Speech and Audio Processing Technical Field Award will be presented to Mark Yoffe Liberman “for pioneering contributions to and continued leadership in robust, replicable, and data-driven speech and language science and engineering.”



Russ Mersereau

The IEEE Fourier Award for Signal Processing will be presented to Russ Mersereau “for sustained technical contributions to multidimensional digital signal processing.”



Jerome John Blair

The IEEE Joseph F. Keithley Award in Instrumentation and Measurement was awarded to Jerome John Blair “for contributions to test procedures for analog-to-digital and digital-to-analog converters and to enhanced-accuracy gamma-ray spectrometry.”



C.-C. Jay Kuo

The IEEE Leon K. Kirchmayer Graduate Teaching Award will be presented to C.-C. Jay Kuo “for inspirational guidance of graduate students and curriculum development in the area of multimedia signal processing.”



Takao Nishitani



John S. Thompson

The IEEE Donald O. Pederson Award in Solid-State Circuits was awarded to Takao Nishitani and John S. Thompson “for pioneering real-time programmable digital signal processor architectures.”

SP

Visit the SPS Award website for more information and links to award-winning papers of this and past years: <http://signalprocessingsociety.org/get-involved/award-recipients>

Learn more about the SPS award program at:

<http://signalprocessingsociety.org/get-involved/awards-submit-award-nomination>

Digital Object Identifier 10.1109/MSP.2017.2673081



Raviraj Adve



Sos Agaian



R. Michael Buehrer



Lap-Pui Chau



Huaiyu Dai



Ricardo De Queiroz



Jose Dias



Pier Luigi Dragotti



Pablo Estevez



James Fowler



Michael Gastpar



Uwe Hanebeck



Julia Hirschberg



Yo-Sung Ho



Tatsuya Kawahara



Takao Kobayashi



Witold Krzymien



Li Li



Xin Li



Yuanqing Li



Qilian Liang



Teng Joon Lim



Wing-Kin Ma

Congratulations!
*SPS Members Elected
IEEE Fellow Class 2017*



Enrico Magli



Aleksandra Mojsilovic



Srikantan Nagarajan



Panos Nasiopoulos



Brett Ninness



Phillip Pace



Bhiksha Raj



Bhuvana Ramabhadran



Gael Richard



Seb Savory



Behzad Shahraray



Steven Smith



Leif Sörnmo



Yannis Stylianou



Guaning Su



Sabine Susstrunk



Nuno Vasconcelos



Xianbin Wang



Zhen Wang



Ying Wu



Shuicheng Yan



Rui Zhang



Ce Zhu

Digital Object Identifier 10.1109/MSP.2017.2657319

Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays

Rusek, F.; Persson, D.; Lau, B.K.; Larsson, E.G.; Marzetta, T.L.; Edfors, O.; Tufvesson, F.

In this article, the authors report that very large multiple-input, multiple-output (MIMO) offered the unique prospect within wireless communication of saving an order of magnitude, or more, in transmit power. This article points out a fundamental challenge of a very large MIMO system design and surveys some related algorithms.

January 2013

OFDM versus Filter Bank Multicarrier

Farhang-Boroujeny, B.

This article addresses the shortcomings of orthogonal frequency-division multiplexing in some applications and shows that a filter bank multicarrier could be a more effective solution.

May 2011

Super-Resolution Image Reconstruction: A Technical Overview

Park, S.C.; Park, M.K.; Kang, M.G.

One promising approach to increasing spatial resolution of an image is to use signal processing techniques to obtain a high-resolution image (or sequence) from observed multiple low-resolution images. This article introduces the concept of super-resolution algorithms to readers who are unfamiliar with this area and to provide the technical review of various existing super-resolution methods for experts.

May 2003

Modulation Formats and Waveforms for 5G Networks: Who Will be the Heir of OFDM? An Overview of Alternative Modulation Schemes for Improved Spectral Efficiency

Banelli, P.; Buzzi, S.; Colavolpe, G.; Modenini, A.; Rusek, F.; Ugolini, A.

This article provides a review of some modulation formats suited to fifth-generation (5G) cellular communications, enriched by a comparative analysis of their performance in a cellular environment, and



by a discussion of their interactions with specific 5G ingredients.

November 2014

Modeling and Optimization for Big Data Analytics: (Statistical) Learning Tools for Our Era of Data Deluge

Slavakis, K.; Giannakis, G.B.; Mateos, G.

This article contributes to the ongoing cross-disciplinary efforts in data science, by putting forth encompassing models, capturing a wide range of signal processing relevant data analytic tasks. It offers scalable architectures and optimization algorithms for decentralized and online learning problems, while revealing fundamental insights into the various analytic and implementation tradeoffs involved. The close connections of the presented framework with several big data tasks are highlighted.

September 2014

Locating the Nodes: Cooperative Localization in Wireless Sensor Networks

Patwari, N.; Ash, J.N.; Kyperountas, S.; Hero, O.; Moses, R.L.; Correal, N.S.

This article provides a window into cooperative localization, which has found considerable applications in ad hoc and wireless sensor networks. It presents measurement-based statistical models of time of arrival, angle of arrival, and received signal strength and uses them to generate localization performance bounds. It also surveys a large and growing body of sensor localization algorithms. This article is intended to emphasize the basic statistical signal processing background, necessary to understand the state-of-the-art methods, and to make progress in the new and largely open areas of sensor network localization research.

July 2005

Big Data Analysis with Signal Processing on Graphs: Representation and Processing of Massive Data Sets With Irregular Structure

Sandryhaila, A.; Moura, J.M.F.

This article reviews fundamental concepts of discrete signal processing on graphs, including graph signals and graph filters, graph Fourier transform, graph frequency, and spectrum ordering, and compares them with their counterparts from the classical signal processing theory. It considers product graphs as a graph model that helps extend the application to large data sets through efficient implementation based on parallelization and vectorization. The article also relates the presented framework to existing methods for large-scale data processing.

September 2014

Communicating While Computing: Distributed Mobile Cloud Computing Over 5G Heterogeneous Networks

Barbarossa, S.; Sardellitti, S.; Lorenzo, P.D.

In this article, the authors proposed a system perspective of the next fifth-generation (5G) systems centered on the need to empower energy-hungry mobile terminals with computation offloading capabilities via proximity to radio access through small-cell base stations, endowed with cloud functionalities. They also showed how the optimal resource allocation involves a joint allocation of radio and computation resources, within a fully cross-layer approach.

November 2014

Speaker Recognition by Machines and Humans: A Tutorial Review

Hansen, J.H.L.; Hasan, T.

In this article, the authors review the literature on speaker recognition by machines and humans, with an emphasis on prominent speaker-modeling techniques that have emerged in the last decade for automatic systems. They conclude this review with a comparative study of human versus machine speaker recognition and attempt to point out strengths and weaknesses of each.

November 2015

(continued on page 122)

SPECIAL REPORTS

John Edwards

Stepping Up Security with Signal Processing

Innovative tools and approaches address threats on multiple fronts

Signal processing is helping to make life a little less scary. Cyberattacks, street crime, terrorist strikes, and other types of public threats have the potential to devastate lives, destroy businesses, and even bring down governments. Finding ways to defuse potential threats, as well as pinpoint perpetrators before they can act, is important for maintaining public security as well as ensuring a better quality of life for people worldwide. According to a recent report issued by the Institute for Economics and Peace, a think tank based in Sydney, Australia, the global economic impact of violence was US\$13.6 trillion in 2015, or 13.3% of the world's gross domestic product.

With the ultimate goal of creating a safer and more stable society, researchers worldwide are developing an array of innovative security tools and approaches that are rooted in or assisted by signal processing. "Security is a wild, wild west frontier where signal processing is one of the tools for explorer," observes Yang Cai, a senior systems scientist at Carnegie Mellon University's CyLab Institute and the director of CyLab's Visual Intelligence Studio.

Visualizing cyberattacks

Distributed denial of service (DDoS) attacks pose a major threat to websites worldwide. A DDoS attack attempts to make an online service unavailable by overwhelming it with traffic from

multiple sources. To help website operators thwart such potentially paralyzing attacks, a CyLab research team, led by Prof. Cai, has developed a novel interactive visual analytics tool that scrutinizes large volumes of network metadata, log files, and security event logs to discover hidden data and suspicious dynamic behavioral patterns lurking within a large network infrastructure (Figure 1).

Visual Analytics for Discovering Dynamic Patterns in Network Data—code-named *Meteor*—enables semantic representations of cyberdynamic compo-

nents, such as packet and protocol sequences, attack behaviors, insider threats, and sequential patterns, as well as access points and cyber events. A three-dimensional (3-D) navigation interface enables users to browse through and analyze a large volume of cyberdata, including network traffic and historical data.

The researchers have already used *Meteor* to study a typical malware distribution network (MDN), a connected set of maliciously compromised domains that's used to disseminate compromised software to unknowing computers and



FIGURE 1. Prof. Yang Cai (left) and research assistant Sebastian Peryt examine a visualization at the Carnegie Mellon University CyLab Visual Intelligence Studio. (Photo courtesy of Carnegie-Mellon University.)

Digital Object Identifier 10.1109/MSP.2016.2638478
Date of publication: 3 March 2017

users. The team defined the graph of an MDN to visualize top-level-domain (TLD) data collected from Google Safe Browsing reports in a temporal manner, characterizing the topological structure. From the collected data, they were able to identify and label a TLD's role in malware distribution. The visual analytics provided insights into the topological structure of MDNs over time, including highly connected and persistent TLDs and subnetworks. "To visualize the time, frequency, and volume of the network events, we introduced a spectrogram of packet traffic," Prof. Cai says. "This can help to illuminate the temporal patterns, similar to music patterns."

Prof. Cai notes that digital signal processing is necessary to filter and preprocess massive amounts of raw data prior to visualization. "It's like finding a needle in a haystack," he says. "We need to reduce the haystack to a manageable size before we can dive in for a detailed search." Signal processing can also be used to characterize dynamic patterns such as periodicity and anomalies, he added.

To process the massive data set, the researchers developed signal processing algorithms designed to filter the data and illuminate potential patterns. "These algorithms include short-term Fourier transform, graph generation, and human visual signal processing speed matching," Prof. Cai says.

Prof. Cai believes that the new visualization method will enable users to discover threat clue patterns, such as missing data, malware spread patterns, and novel hacking tactics. He points to attack behavior profiling, inside threat analysis, situation awareness, malware spread pattern discovery, and real-time anomaly detection as potential Meteor applications. The system can be integrated into existing platforms, such as data cloud virtual machines. It can also work with cyberphysical systems, such as the Internet of Things, in the context of data anonymity, user authentication, and phishing defense.

Addressing data set complexity was the project's biggest signal processing challenge. "To visualize the dynamics of evolving patterns, we needed to periodically sample the data with time stamps,"

Prof. Cai says. In many cases, however, the data are not perfectly collected. "There are missing data points or overlaid data—even fake data," he continues. "Furthermore, many network events and anomalies are not repeated: the 'Black Swan' phenomena (an unpredictable or unforeseen event)."

"Interactive visual analytics is a learning process," Prof. Cai says. "After we learn meaningful visual patterns, we want to model the visual ontology computationally," he notes. "We anticipate that digital signal processing will be a powerful tool to represent features and patterns in the data." The researchers also plan to study the link between visual analytics and machine learning. "Once we learn the visual patterns, we want to feed them into a machine that is able to learn automatically," he states. "After all, we want to develop a human-computer interaction system that combines human visual intelligence with machine intelligence to maximize the efficiency of data analyses."

High-resolution facial detection

As video resolution increases, existing facial detection technologies struggle to perform rapid and accurate analytics on high-definition (HD) streams. "At present, there is no shortage in the resolution of cameras," says Ilya Kalinovskii, a doctoral student at Tomsk Polytechnic University in Tomsk, Russia, who has developed an algorithm that's designed to bring fast and reliable facial detection to a new generation of HD surveillance systems. "Already available in the market are cameras that can transmit an 8K video stream," he said.

The challenge HD video presents to existing video analytics technologies is their need for very large computational resources and high-bandwidth data channels. "My algorithm allows solving this problem, at least in biometric video analytics systems," Kalinovskii says, noting that the technology will enable high-resolution automatic monitoring of a local

environment, even potentially allowing emergency situations to be detected quickly without operator attention.

Existing video-based facial detection systems often rely on the Viola-Jones algorithm framework, which until recently has allowed reliably high detection rates. "However, its performance is not enough to process 4K videos in real time; it also allows a great number of false positive detections," says Kalinovskii, who believes that his new facial detection method marks a

significant improvement to Viola-Jones. "It provides greater accuracy in face detection and up to a ten times higher rate of the frame processing," he says. "It is the first algorithm capable to process 4K-video in

real-time mode on low-power computing devices" (Figure 2).

The new technology takes advantage of sophisticated deep learning methods. "A face detector is based on a compact convolutional neural network (CNN) cascade," Kalinovskii says. Deep artificial neural networks, particularly a CNN, are rapidly gaining popularity, largely due to the fact that they can outperform human accuracy in solving a number of video image detection tasks. "Traditionally, however, to achieve a high quality of classification, one needed to use very large neural networks with millions of parameters that require powerful and expensive graphics processing units to work," Kalinovskii says.

Working with Prof. Vladimir Spitsyn of the school's Computer Engineering Department, Kalinovskii developed CNN models requiring only a few thousand parameters. "This determined the exceptional performance of the detector, which is capable of operating in real time, even on a central processing unit," Kalinovskii says.

Optimizing CNN computations posed the project's biggest challenge. "The CNN cascade itself does not provide the needed performance," Kalinovskii says. "Only the combination of efficient algorithms for image filtering and computations of

Existing video-based facial detection systems often rely on the Viola-Jones algorithm framework, which, until recently, has allowed reliably high detection rates.

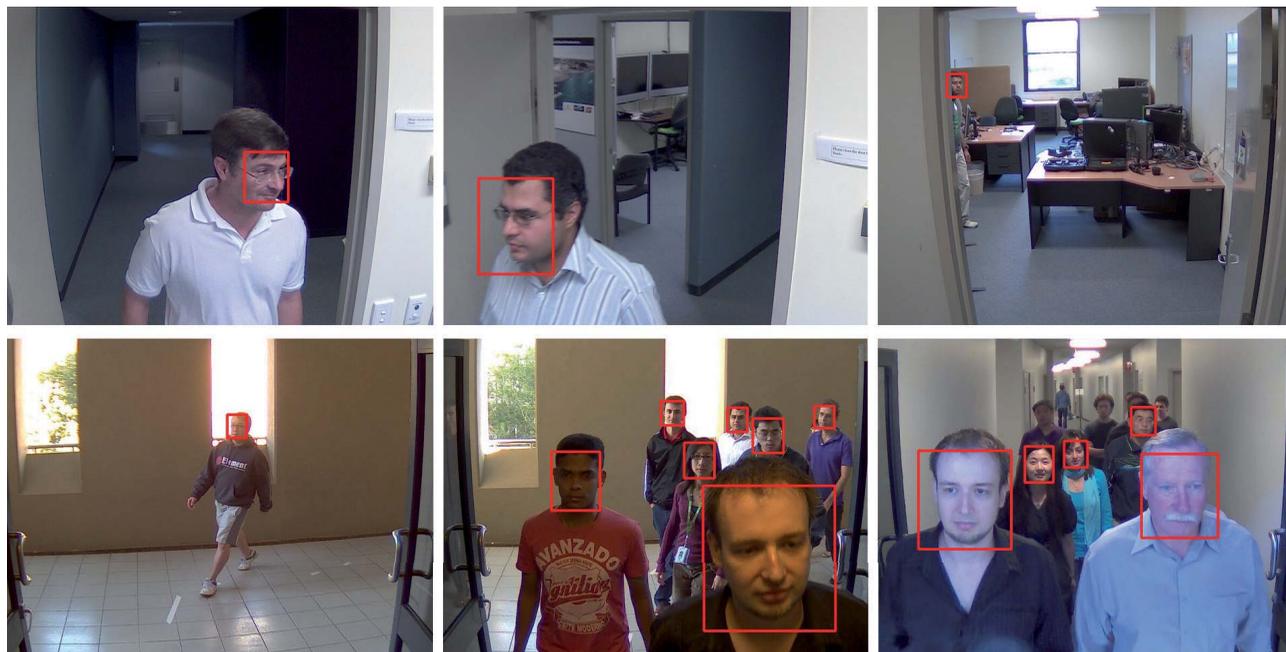


FIGURE 2. A new algorithm developed by a Tomsk Polytechnic University researcher promises improved facial detection in high-resolution video streams. (Photos courtesy of Tomsk Polytechnic University.)

nonlinear transformations has allowed achieving the high rate of data processing that made it possible to process HD video in real time.” CNN training, meanwhile, was a necessary but time consuming process. “Training neural networks is a sophisticated problem and required fine-tuning of lot of hyperparameters,” Kalinovskii says.

Kalinovskii feels that there will be a growing demand for the type of technology he has developed. “To detect a person by his facial image in a crowd, like in a stadium or subway, it is required to work with an image in a scene with very high detail,” he says. “4K resolution is suitable for this purpose.”

Looking to the future, Kalinovskii hopes to bring the algorithm to Android and iOS devices, which could make high-resolution facial detection a mobile technology that could be used almost anywhere. “Currently I am working on porting the algorithm onto advanced RISC machine processors, which require more code optimization to improve performance,” he says.

Protecting vulnerable grids

Electrical power grids worldwide are aging, overburdened, and susceptible

to attack and failure at any moment, potentially leaving millions of people, as well as essential public services, literally in the dark. With the goal of improving grid security and reliability, Binghamton University, New York, researchers believe that a singular spectrum analysis (SSA) algorithm could become the best tool to help authorities remotely detect and locate power grid faults.

A power grid—a collection of multiple complete circuits—is designed to keep electricity flowing even when an individual circuit fails. Although a power grid’s inherent redundancy creates a great deal of stability, the technology is also highly complex and populated with vulnerable points. By leveraging a power grid’s weak spots, an attacker can break in and subtly change electricity flow, creating a cascading and potentially catastrophic effect on the entire infrastructure.

“The SSA algorithm enables quick and accurate detection of subtle changes in the sampled signal series without any knowledge of ‘normal’ or ‘abnormal’ statistical features,” says research team leader Prof. Yu Chen, a Binghamton University associate professor of electrical and computer engineering. “It is an ideal match

to power grid monitoring tasks, where it is very challenging to have a comprehensive model of behavior patterns.” The researchers call their approach *model-free online distributed disturbance location* (Figure 3).

As things currently stand, any time or location anomalies lurking within a grid—potential signs of a hidden problem—are determined by familiar formulas, such as the event start time (EST) algorithm, which calculates the arrival times of power changes at different geographic locations. Although the arrival differences are very small, they generally provide sufficient information to triangulate a location. The Binghamton researchers used simulation data to show that the SSA algorithm can be faster and more effective than EST at finding anomalies in the power grid created by either generator or transmission line issues. According to Prof. Chen, SSA works better in noisy environments and is able to capture subtle disturbances that elude the EST algorithm. “Wide-area dynamic events are currently not well presented to operators,” Prof. Chen says. Project coresearchers are Prof. Aleksey Polunchenko from the Department of Mathematical Sciences, Prof. Ning Zhou

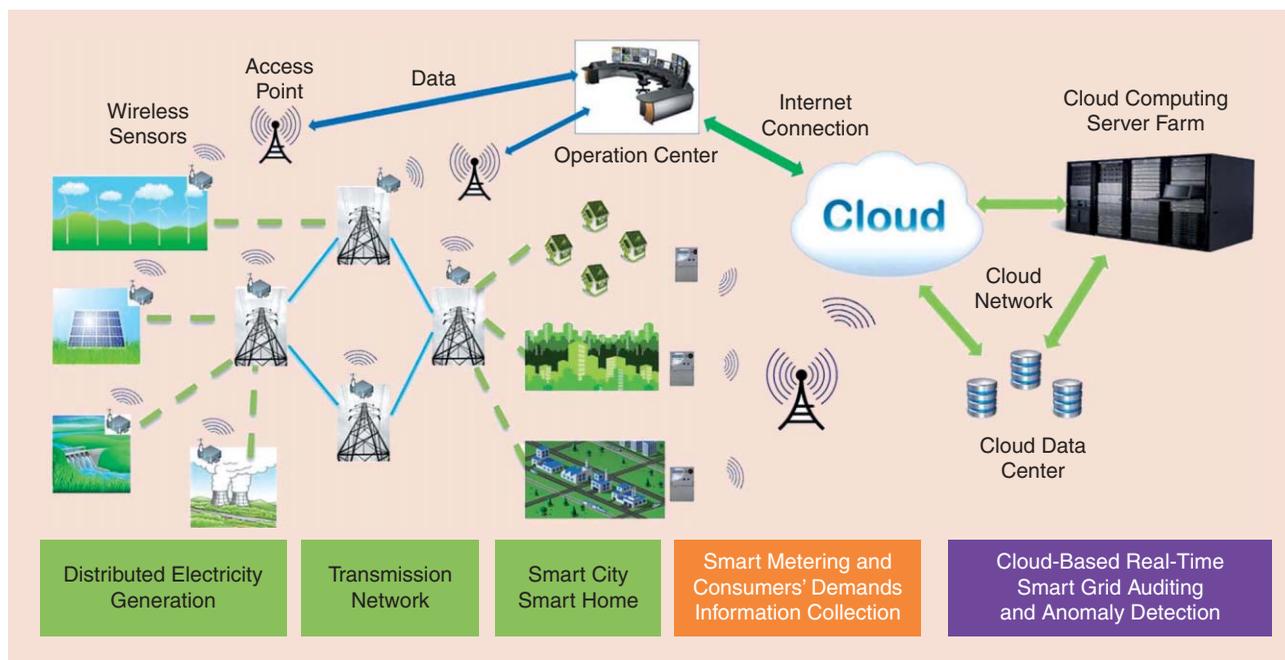


FIGURE 3. Binghamton University researchers have proposed a new SSA algorithm-based model for remotely detecting and locating power grid faults. This diagram shows how a cloud-based monitoring system using the model might be configured. (Image courtesy of Binghamton University.)

from the Department of Electrical and Computer Engineering, and doctoral student Zekun Yan from the Department of Electrical and Computer Engineering.

The SSA-based change point detection method is a nonmodel and structure-free data analysis technique, Prof. Chen says. “As a result, on one hand, theoretically, it can detect faults of any category because SSA don’t have a limitation for data structure.” On the other hand, Prof. Chen notes, history records and prior information about the fault condition are not needed to build the model. “In addition, experimental results have shown high accuracy for this method,” Prof. Chen says.

At the individual local sensor level, the moment a disturbance arrives at the measured position is a critical factor in disturbance detection. “With the help of synchronization, the disturbance record data are marked by time instance such that it can be compared with other records,” Prof. Chen says. “Therefore, naturally, it’s a time series and suitable for a corresponding data analysis technique.”

The new approach utilizes a traveling wave-based approach that’s both simple and well suited for use in the wide-area systems, such as power grids. “The trav-

eling wave-based method focuses on the waves generated by the disturbance and the correlation between the forward and backward wave arrival or direct detection of the arrival time on a single line,” Prof. Chen says. Specifically, the approach uses a time difference of arrival (TDOA) method. “The TDOA method assumes a unified wave traveling velocity and utilizes the different arrival times at different sensors to locate the wave start point—the fault source,” Prof. Chen says. “The critical challenges for the method are the arrival time detection accuracy and the wave traveling velocity error.”

Developing a highly accurate power grid detection/location technology poses some unique challenges. “First, the propagation speeds of the electromechanical signals vary from 100 miles to 1,000 miles per second among different paths in the power grid due to the adoption of variant transmission technologies,” Prof. Chen says. “Second, absolute geological distance is used to simplify the calculation in spite of the fact that the electromechanical waves propagate along with the power grid topology.”

Although the individual measurement data has propagation error and modifica-

tion error, a proposed temporal scanning algorithm would take advantage of the abundant amount of information supplied by frequency disturbance recorders dispersed across the grid to reach an accurate value. “According to the law of large number, with the large number of samples the estimate will gradually converge to the true value,” Prof. Chen says.

The researchers continue to refine the system and are currently addressing two key issues: the fast detection of weak oscillations and the prediction of potential vibrations based on the detection of disturbances and oscillations.

Meanwhile, the current SSA algorithm can only detect and locate problems; it cannot predict what may happen in the future. The researchers hope that the current platform’s ability to detect subtle power grid change changes could someday be improved to the point where it will be able to forecast problems before they occur.

Author

John Edwards (jedwards@johnedwardsmedia.com) is a technology writer based in the Phoenix, Arizona, area.

SP

FROM THE GUEST EDITORS

John H.L. Hansen, Kazuya Takeda, Sanjeev M. Naik, Mohan M. Trivedi, Gerhard U. Schmidt, Yingying (Jennifer) Chen, and Wade Trappe

Signal Processing for Smart Vehicle Technologies: Part 2

The automobile has played an important role in shaping our society for the past century, and signal processing is playing an increasingly substantial role in shaping smart vehicle technologies today. Part 2 of the special section on signal processing for smart vehicle technologies provides additional six articles, complementing those in Part 1, which was published in the November 2016 issue of *IEEE Signal Processing Magazine (SPM)* [1].

Digital Object Identifier 10.1109/MSP.2017.2650299
Date of publication: 3 March 2017

Part 2 starts with two articles that address radar-based sensing technologies.

The first article, “Automotive Radars,” by Patole et al., provides a tutorial overview on signal processing techniques used in radar sensing in automobiles. In “Advances in Automotive Radar,” Engels et al. focus on computationally efficient algorithms to support high-resolution frequency estimation in automotive radar applications.

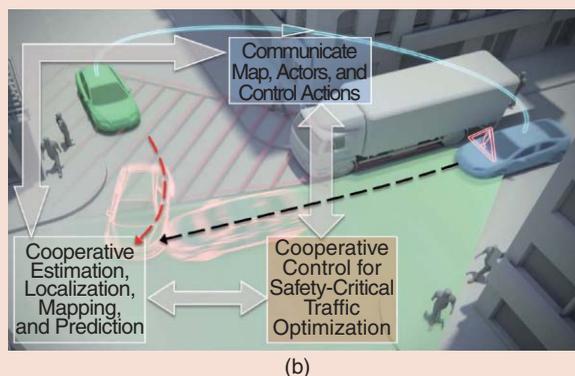
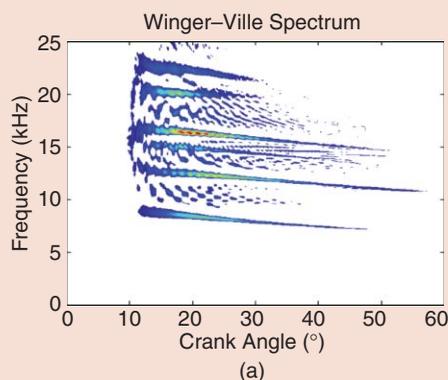
Given the importance of communication capabilities and location information for users on the

move, Schwarz et al. surveyed the related efforts and developments of the fourth-generation (4G) wireless communications in their article “Signal Processing Challenges in Cellular-Assisted Vehicular Communi-

cations.” Additionally, “The Future of Automotive Localization Algorithms” by Karlsson and Gustafsson addressed the

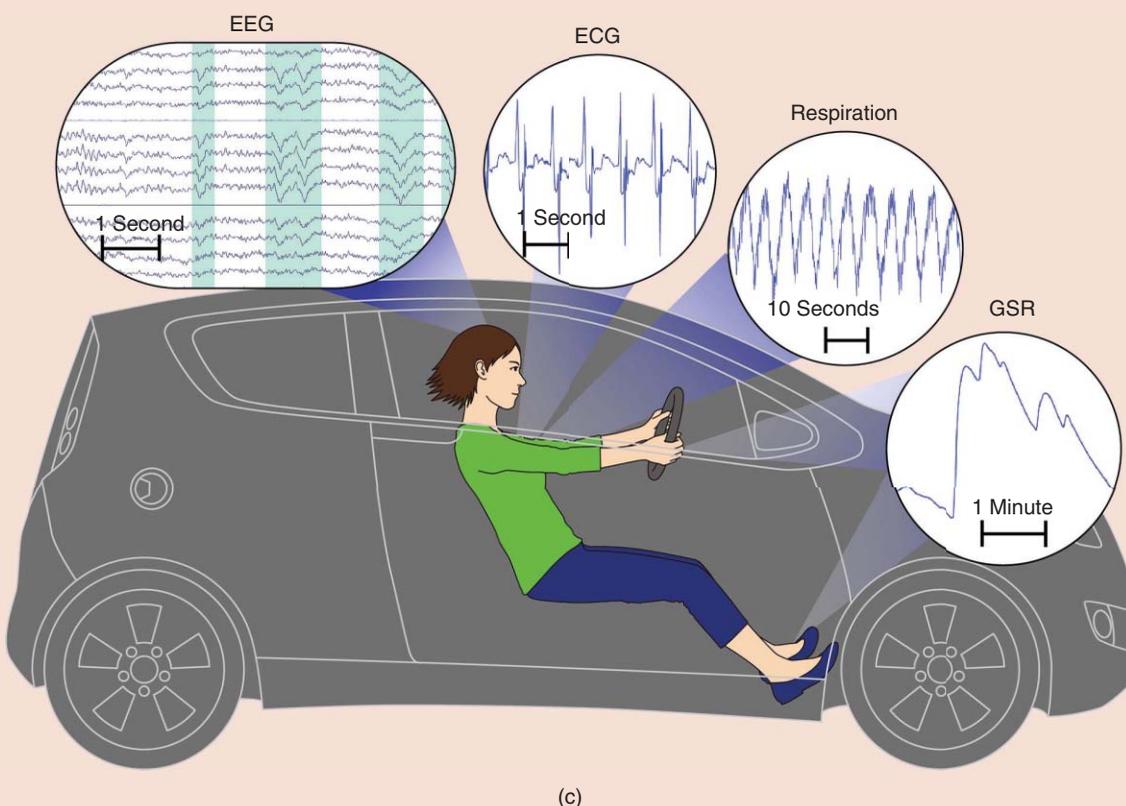
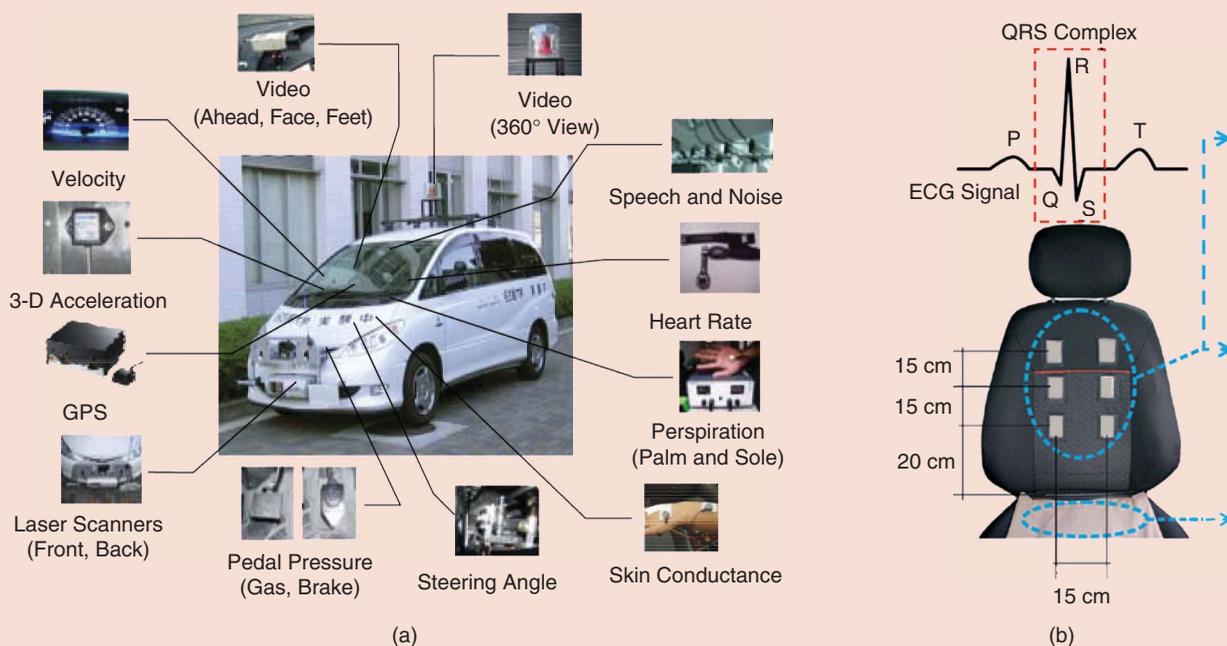
As a picture is worth 1,000 words, this guest editorial includes a visual summary consisting of figures selected from articles in both parts in this two-part special issue.

Signal Processing Supporting Vehicular Control and Coordination



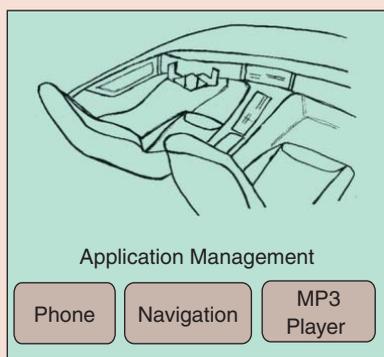
Signal processing provides essential support in vehicular control and coordination: (a) digital signal processing (DSP) applications in engine control and onboard diagnostics (see the article by Wu and Naik, on page 70 of this issue), and (b) coordination of cooperative autonomous vehicles (see the article by Hult et al., page 74 in *SPM*'s November 2016 issue).

Driver Monitoring

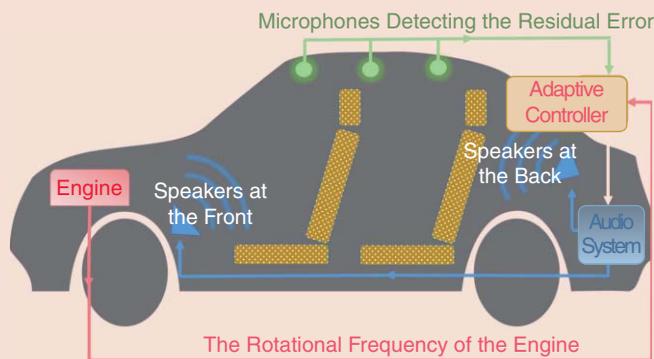


A variety of signal sensing and analytic techniques are being developed to provide driver status monitoring to enhance the safety of driving: (a) driver-behavior modeling using on-road driving data (see the article by Miyajima and Takeda, page 14 in *SPM's* November 2016 issue), (b) driver status monitoring systems for smart vehicles using physiological sensors (see the article by Choi et al., page 22 in *SPM's* November 2016 issue), and (c) smart driver monitoring where signal processing meets human factors (see the article by Aghaei et al., page 35 in *SPM's* November 2016 issue).

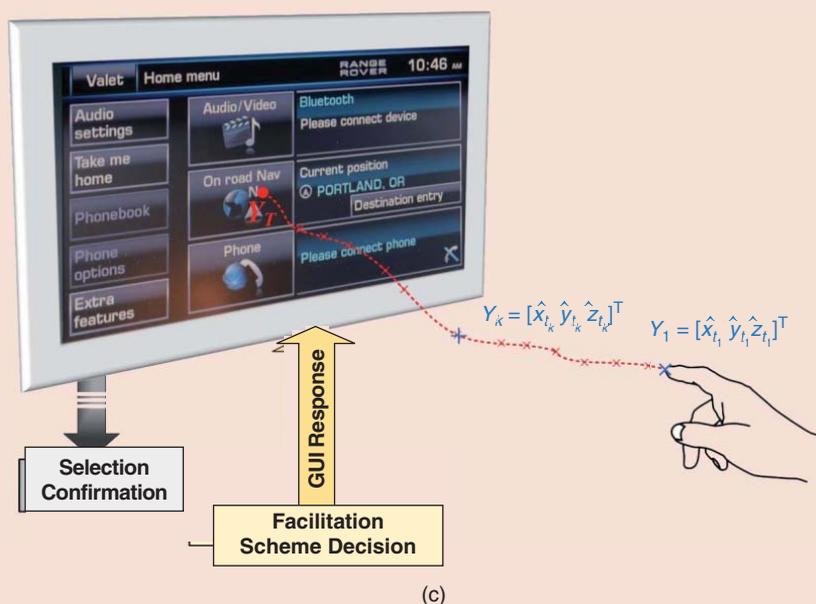
Enhanced User Interface in Automobiles



(a)



(b)



(c)

Signal processing techniques are essential in enhancing the user interface in an automobile: (a) conversational in-vehicle dialog systems (see the article by Weng et al., page 49 in *SPM's* November 2016 issue), (b) active noise control inside automobile cabins (see the article by Samarasinghe et al., page 61 in *SPM's* November 2016 issue), and (c) intelligent interactive displays in vehicles with intent prediction (see the article by Ahmad et al., page 82 of this issue).

technologies behind making localization available, reliable, and scalable anywhere and at any time.

Signal processing is also in the heart of keeping automobiles running and enabling greener automobiles. These aspects are surveyed in the article “DSP Applications in Engine Control and Onboard Diagnostics” by Wu and Naik. With the increasing adoption of interactive displays in the vehicle, Ahmad et al. addressed the unique issues and a

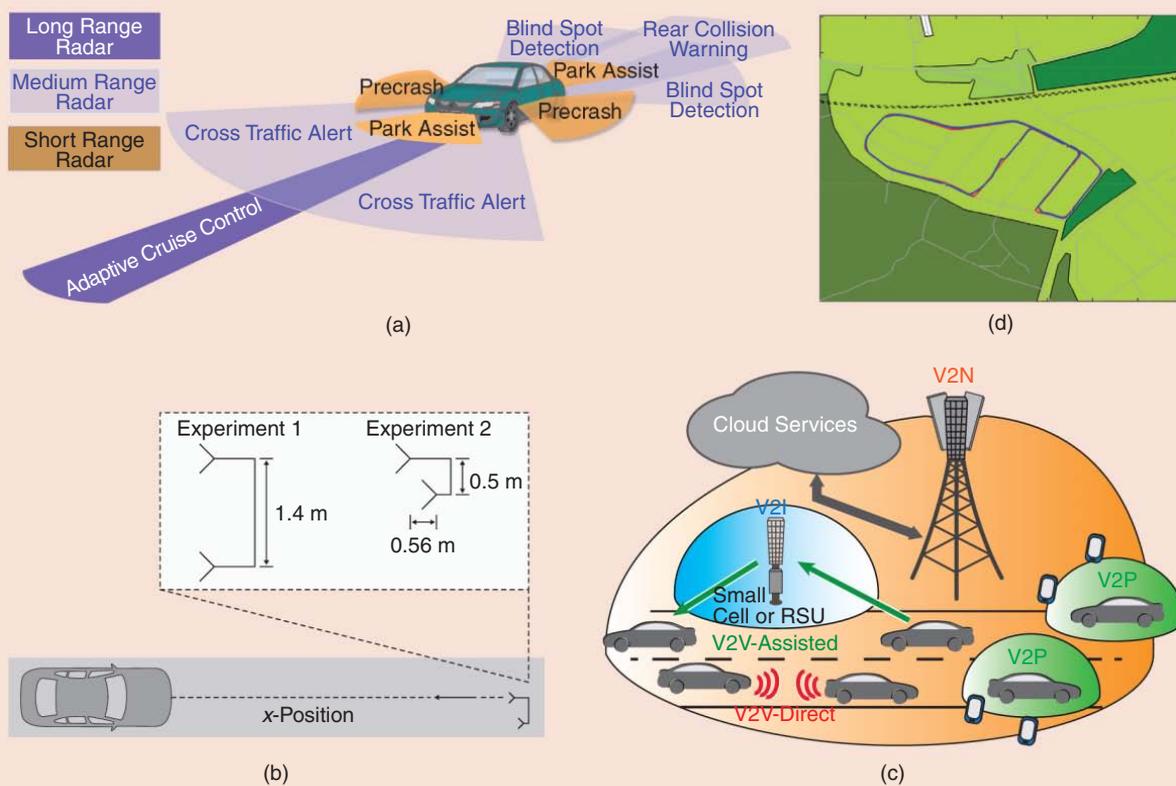
Bayesian framework in this user-interface setting in the article “Intelligent Interactive Displays in Vehicles with Intent Prediction.”

As a picture is worth 1,000 words, this guest editorial includes a visual summary consisting of figures selected from articles in both parts in this two-part special issue: see “Signal Processing Supporting Vehicular Control and Coordination,” “Driver Monitoring,” “Enhanced User Interface

in Automobiles,” and “Processing in Automotive Radar Sensing, Communications, and Localization.” Thanks are given to *SPM* Editor-in-Chief Dr. Min Wu and IEEE Signal Processing Society Publications Administrator Rebecca Wollman for their support, as well as to the numerous anonymous reviewers.

Once again, we welcome you to explore these articles, as well as the field of signal processing for vehicular

Processing in Automotive Radar Sensing, Communications, and Localization



Signal processing techniques are used in automotive sensing, communications, and localization: (a) an overview of automotive radars (see the article by Patole et al., page 22 of this issue), (b) computationally efficient high-resolution frequency estimation in automotive radars (see the article by Engels et al., page 36 of this issue), (c) cellular-assisted vehicular communications (see the article by Schwarz et al., page 47 of this issue), and (d) automotive localization algorithms (see the article by Karlsson et al., page 60 of this issue).

technologies. Happy reading—and happy driving!

Meet the Guest Editors

John H.L. Hansen (john.hansen@utdallas.edu) is the associate dean for research and a professor of electrical engineering at the University of Texas at Dallas. He served as lead guest editor for this special section on smart vehicle technologies.

Kazuya Takeda (kazuya.takeda@nagoya-u.jp) is a professor at the Graduate School of Informatics and Green Mobility Collaborative Research Center, Nagoya University, Japan.

Sanjeev M. Naik (sanjeev.m.naik@gm.com) is the engineering group manager in advanced controls engineering at the General Motors.

Mohan M. Trivedi (mtrivedi@ucsd.edu) is a distinguished professor of electrical and computer engineering at the University of California, San Diego.

Gerhard U. Schmidt (gus@tf.uni-kiel.de) is a full professor with Kiel University, Germany.

Yingying (Jennifer) Chen (yingying.chen@stevens.edu) is a professor of electrical and computer engineering at Stevens Institute of Technology in Hoboken, New Jersey.

Wade Trappe (trappe@winlab.rutgers.edu) is a professor of electrical and computer engineering at Rutgers University. He served as area editor for special issues of *IEEE Signal Processing Magazine* during the effort of this special section on smart vehicle technologies.

Reference

[1] J. H. L. Hansen, K. Takeda, S. M. Naik, M. M. Trivedi, G. U. Schmidt, and Y. Chen, "Signal Processing for Smart Vehicle Technologies," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 12–13, Nov. 2016.

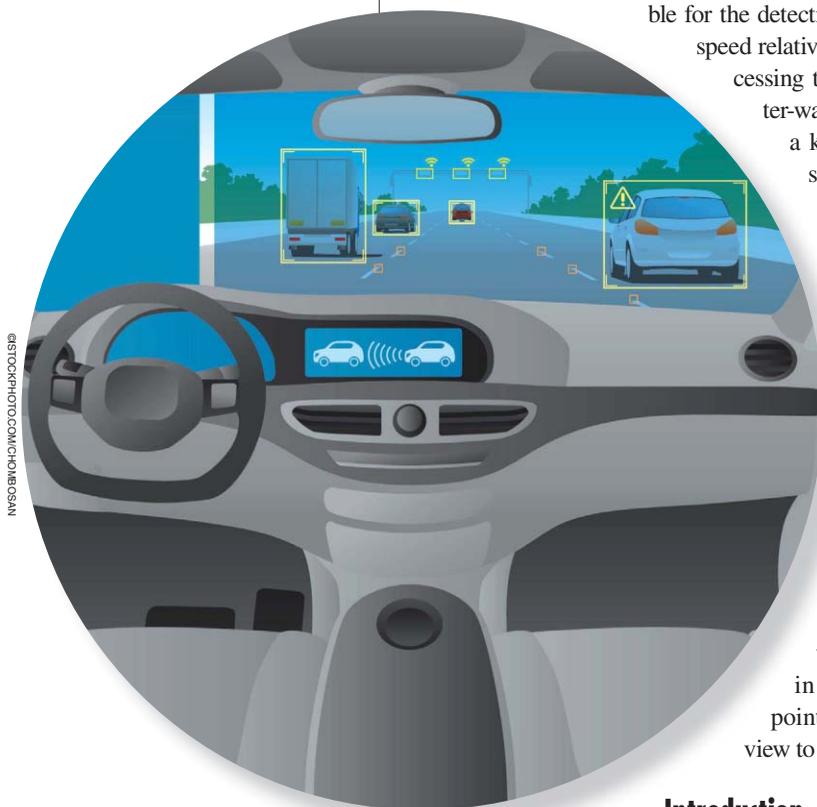


Sujeet Patole, Murat Torlak, Dan Wang, and Murtaza Ali

Automotive Radars

A review of signal processing techniques

Automotive radars, along with other sensors such as lidar, (which stands for “light detection and ranging”), ultrasound, and cameras, form the backbone of self-driving cars and advanced driver assistant systems (ADASs). These technological advancements are enabled by extremely complex systems with a long signal processing path from radars/sensors to the controller. Automotive radar systems are responsible for the detection of objects and obstacles, their position, and speed relative to the vehicle. The development of signal processing techniques along with progress in the millimeter-wave (mm-wave) semiconductor technology plays a key role in automotive radar systems. Various signal processing techniques have been developed to provide better resolution and estimation performance in all measurement dimensions: range, azimuth-elevation angles, and velocity of the targets surrounding the vehicles. This article summarizes various aspects of automotive radar signal processing techniques, including waveform design, possible radar architectures, estimation algorithms, implementation complexity-resolution trade off, and adaptive processing for complex environments, as well as unique problems associated with automotive radars such as pedestrian detection. We believe that this review article will combine the several contributions scattered in the literature to serve as a primary starting point to new researchers and to give a bird’s-eye view to the existing research community.



Introduction

The history of radio detection and ranging, more commonly known as *radar*, starts with the experiments carried out by Hertz and Hülsmeyer on the reflections of electromagnetic (EM) waves and ideas advocated by Tesla and Marconi in the late 19th and early 20th centuries. Earlier developments in radar technology were limited to military applications such as aircraft/ship surveillance, navigation, and weapons guidance. Radar is now used in many applications, including civilian aviation, navigation, mapping, meteorology, radio astronomy, and medicine. The main

Digital Object Identifier 10.1109/MSP.2016.2628914
Date of publication: 3 March 2017

objectives of a radar system are to detect the presence of one or more targets of interest and estimate their range, angle, and motions relative to the radar [1].

To the everyday person, tangible applications of radar include speed guns used by law enforcement officers to detect speeding drivers. Action heroes in movies sometimes drive a fancy car with attractive features that can track an enemy's speed and location, move swiftly and automatically amid obstacles, and debut its night vision feature during the movie's climax. The ambition of having all of these add-ons to a car has become feasible with the flourishing mm-wave circuit technology and advanced signal processing techniques. Advances in circuit technology reinforced by new signal processing algorithms, machine learning, artificial intelligence, and computer-vision techniques have made self-driving cars a reality.

Such cars also rely on different sensors such as a laser, a camera, ultrasound, global positioning system, and radar. Among these sensors, radar offers the possibility of seeing long distances ahead of the car in poor visibility conditions, which can help avoid collisions [2]. For example, Google's self-driving car [3] has radars mounted on both front and rear bumpers of the vehicle to detect objects in its surroundings.

Automotive radars were first deployed several decades ago. The evolution of automotive radar from its inception to the present has been thoroughly discussed in [4]. With highly integrated and inexpensive mm-wave circuits implemented in silicon, compact automotive radar safety systems have become a popular feature [5], [6]. Since then, review articles written on automotive radar mostly covered the circuit implementation, market analysis, and architectural-level signal processing [7]–[9]. However, there are many aspects of automotive radar signal processing techniques scattered throughout the literature. For example, a part of the literature may concentrate on detecting the presence or absence of

Advances in circuit technology reinforced by new signal processing algorithms, machine learning, artificial intelligence, and computer-vision techniques have made self-driving cars a reality.

targets, while another might look at radar estimation problems concerning their location and velocity in space relative to the radar [10], [11].

This article's goal is to review principal developments in signal processing techniques applied to estimating significant target parameters such as range, velocity, and direction. The article also discusses the characterization of radar waveforms and advanced estimation techniques that enhance the operation of automotive radars. In particular, we review each topic with adequate mathematical

framework so as to make this a good start-up document for the newcomer in the field.

Automotive radar classification

Both autonomous and human-driven cars are increasingly using radars to improve drivers' comfort and safety. For instance, park assist and adaptive cruise control provide comfort, while warning the driver of imminent collisions and overriding control of the vehicle to avoid accidents improve the safety. Figure 1 depicts various such radar subsystems that form ADASs. Each subsystem has unique functionality and specific requirements in terms of radar range and angular measurement capability (Table 1). The next section explains the fundamentals of location and speed estimation using the radar measurements.

Basic automotive radar estimation problems

A radar can simultaneously transmit and receive EM waves in frequency bands ranging from 3 MHz to 300 GHz. It is designed to extract information [i.e., location, range, velocity and radar cross section (RCS)] about targets using the EM waves reflected from those targets. Automotive radar systems typically operate at bands in 24 GHz and 77 GHz portions of the EM spectrum known as mm-wave frequencies so that higher velocity and range resolution can be achieved. Fundamental

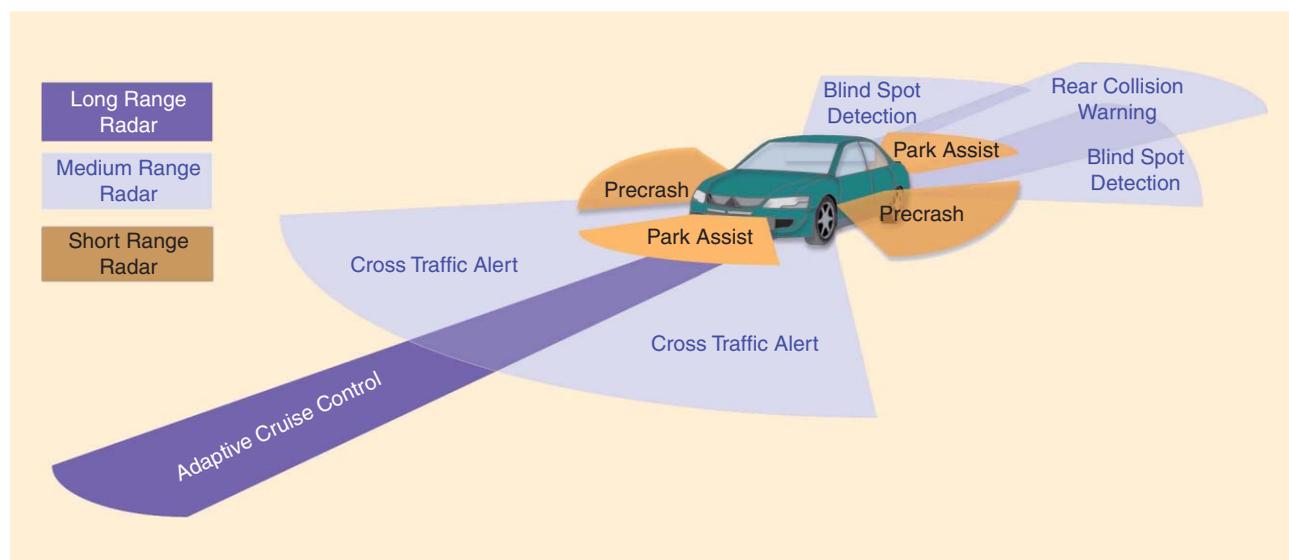


FIGURE 1. An ADAS consists of different range radars.

Table 1. The classification of automotive radars based on range measurement capability.

Radar Type	Long-Range Radars	Medium-Range Radars	Short-Range Radars
Range (m)	10–250	1–100	0.15–30
Azimuthal field of view (deg.)	± 15°	± 40°	± 80°
Elevation field of view (deg.)	± 5°	± 5°	± 10°
Applications	Automotive cruise control	Lane-change assist, cross-traffic alert, blind-spot detection, rear-collision warning	Park assist, obstacle detection, precrash

Classification can be made based on the operating frequency into 24–29 GHz and 76–81-GHz bands [12].

radar operation involves three main tasks: range (distance), relative velocity, and direction estimation, as discussed next.

Range estimation

The range estimation is fundamental to automotive radars. The range R , to a target, is determined based on the round-trip time delay that the EM waves take to propagate to and from that target: $R = (c\tau/2)$, where τ is the round-trip time delay in seconds and c is the speed of light in meters per second ($c \approx 3 \times 10^8$ m/s). Thus, the estimation of τ enables the range measurement [1]. The form of the EM waves (signals) that a radar transmits is important for round-trip time delay estimation. For example, pulse-modulated continuous waves (CWs) consist of periodic and short power pulses and silent periods. Silent periods allow the radar to receive the reflected signals and serve as timing marks for radar to perform range estimation as illustrated in Figure 2. However, unmodulated CW signals (i.e., $\cos(2\pi f_c t)$) cannot be used for range estimation since they lack such timing marks. Additionally, the signal reflected from a target should arrive before the next pulse starts. Hence, the maximum detectable range of a radar depends on pulse repetition interval T_{PRF} . The transmitted signal from the radar until it is received back undergoes attenua-

tion due to the path loss and imperfect reflection from the target. In addition, received target signals are subject to internal noise in radar electronics and interference that may be a result of reflected signals from objects not of interest and may come from human-made sources (i.e., jamming). The typical round-trip time delay estimation problem considers only ambient noise in the form of additive white Gaussian random process. It is assumed that demodulation has already removed the carrier so that a target signal $x(t)$ at baseband can be modeled as

$$x(t) = \alpha s(t - \tau) + w(t), \tag{1}$$

where α is a complex scalar whose magnitude represents attenuation due to antenna gain, path loss, and the RCS of the target and $w(t)$ is additive white Gaussian noise with zero mean and variance σ^2 . The goal is to estimate τ with the complete knowledge of the transmitted radar waveform $s(t)$. Assuming the signal $s(t)$ has unit amplitude and finite energy E_s , the ideal radar receiver can be found using a matched filter with the impulse response $h(t) = s^*(t - \tau)$, which maximizes signal to noise ratio ($\text{SNR} = (\alpha^2 E_s / \sigma^2) = (\alpha^2 T_p / \sigma^2)$) at the output. Thus, the matched filter-based receiver finds the correlation between the transmitted signal and received reflected pulses

$$y(\tau) = \int x(t) s^*(t - \tau) dt. \tag{2}$$

The maximum likelihood (ML) estimate of the time delay is the time that the magnitude of the matched filter output peaks at

$$\hat{\tau} = \arg \max_{\tau} |y(\tau)|. \tag{3}$$

The presence of the noise can perturb the location of the peak, which will result in the estimation error. Furthermore, the radar needs to decide whether or not a received signal actually contains an echo signal from a target. A good deal of classical radar literature is devoted to developing strategies that provide the most favorable detection performance.

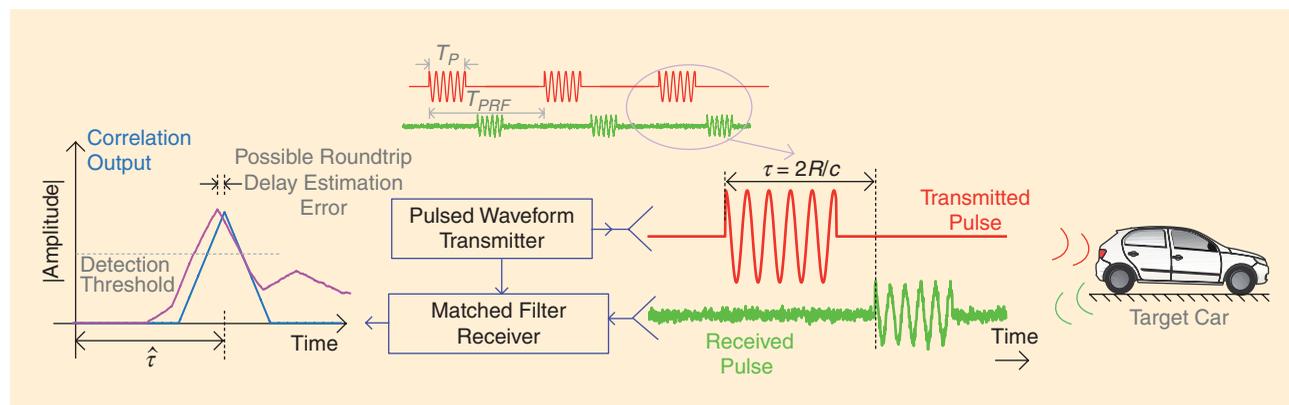


FIGURE 2. A pulsed CW radar with a correlation-based receiver can measure range R of the target car.

A typical decision strategy can be formulated based on statistical hypothesis testing (a target present or not). This leads to a simple threshold testing at the matched filter output.

Range resolution, another key performance measure, denotes the ability to distinguish closely spaced targets. Two targets can be separated in the range domain only if they produce nonoverlapping returns in the time domain. Hence, the range resolution is proportional to the pulsewidth T_p . In other words, finer pulses provide higher resolution. However, shorter pulses contain less energy, which implies poor receiver signal-to-noise ratio (SNR) and detection performance. As explained in the section “Radar Waveforms,” this problem is overcome by the technique called *pulse compression*, which uses phase or frequency modulated pulses.

Velocity estimation

Estimation of the target velocity is based on the phenomenon called the *Doppler effect*. Suppose the car displayed in Figure 2 is moving ahead with differential velocity v . With the existence of relative motion between two cars, the reflected waves are delayed by time $\tau = (2(R \pm vt)/c)$. The time dependent delay term causes a frequency shift in the received wave known as the *Doppler shift* $f_d = (\pm 2v/\lambda)$. The Doppler shift is inversely proportional to wavelength λ , and its sign is positive or negative, depending on whether the target is approaching or moving away from the radar. While this frequency shift can be detected using CW radar, it lacks the ability to measure the targets range. Here, we discuss a pulsed radar configuration that uses frequency modulated (FM) CW pulses and provides simultaneous range-velocity estimation in multitarget traffic scenarios.

The FMCW radar transmits periodic wideband FM pulses, whose angular frequency increases linearly during the pulse. For the carrier frequency f_c and FM modulation constant K , a single FMCW pulse can be written as [see Figure 3(a)]

$$s(t) = e^{j2\pi(f_c + 0.5Kt)t} \quad 0 \leq t \leq T. \quad (4)$$

The signal reflected from a target is conjugately mixed with the transmitted signal to produce a low-frequency beat signal, whose frequency gives the range of the target. This operation is repeated for P consecutive pulses. Two-dimensional (2-D) waveforms in Figure 3(c) depict successive reflected pulses arranged across two time indices. The slow time index p simply corresponds to pulse number. On the other hand, the fast time index n assumes that for each pulse, the corresponding continuous beat signal is sampled with frequency f_s to collect N samples within the time duration T . Assuming single target and neglecting reflected signal distortions, the FMCW radar receiver output as a function of these two time indices is given by

$$d(n,p) \approx \exp\left\{j2\pi\left[\left(\frac{2KR}{c} + f_d\right)\frac{n}{f_s} + f_d p T_0 + \frac{2f_c R}{c}\right]\right\} + \omega(n,p). \quad (5)$$

Therefore, as illustrated in Figure 3(c), discrete Fourier transform across fast time n can be applied to obtain beat frequency $f_b = (2KR/c)$ coupled with Doppler frequency f_d . This operation is also known as the *range transform* or *range gating*, which allows the estimation of Doppler shift corresponding to

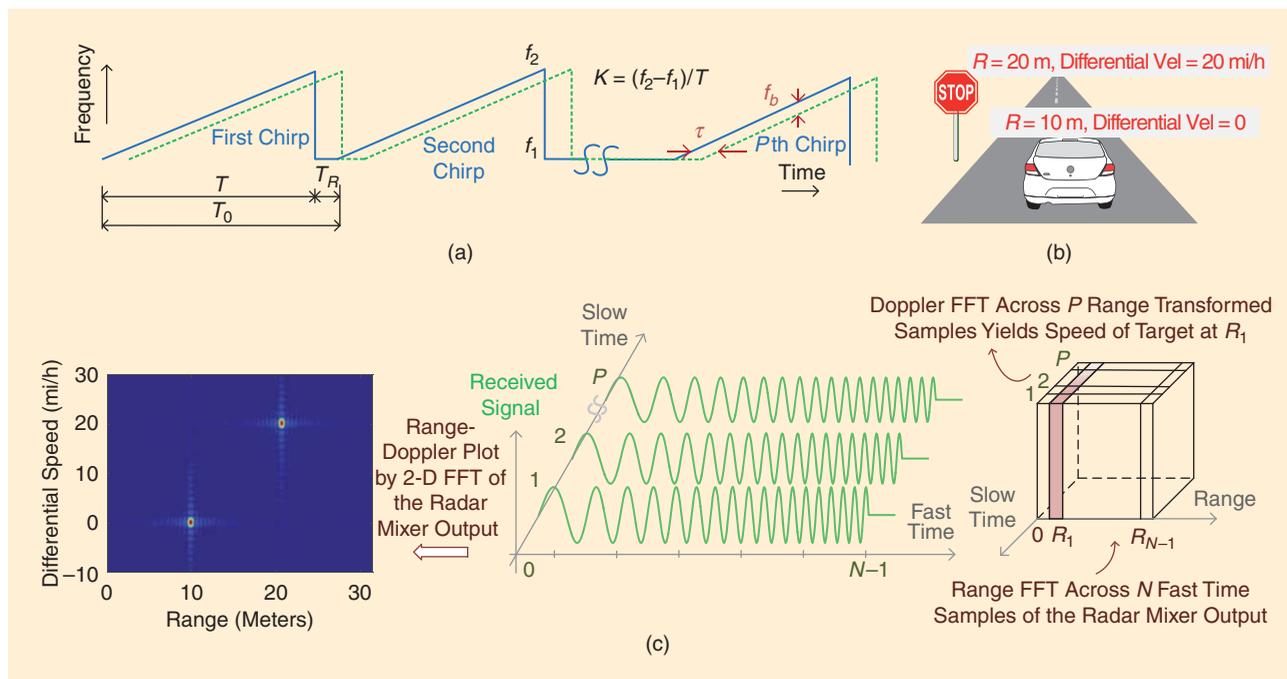


FIGURE 3. (a) A spectrogram of an FMCW waveform with modulation constant $K = (B/T)$, reset time T_R , and pulse period T_b ; transmitting P successive chirps. Round-trip delay τ is converted to beat frequency f_b . (b) Typical traffic scenario: stationary traffic sign, the radar, and passenger car moves at 20 mi/h (range and differential velocity are displayed). (c) A 2-D joint range-Doppler estimation with 77-GHz FMCW radar ($[N, P] = [64, 64]$, SNR = 10 dB, BW = 300 MHz, $T = 300 \mu s$).

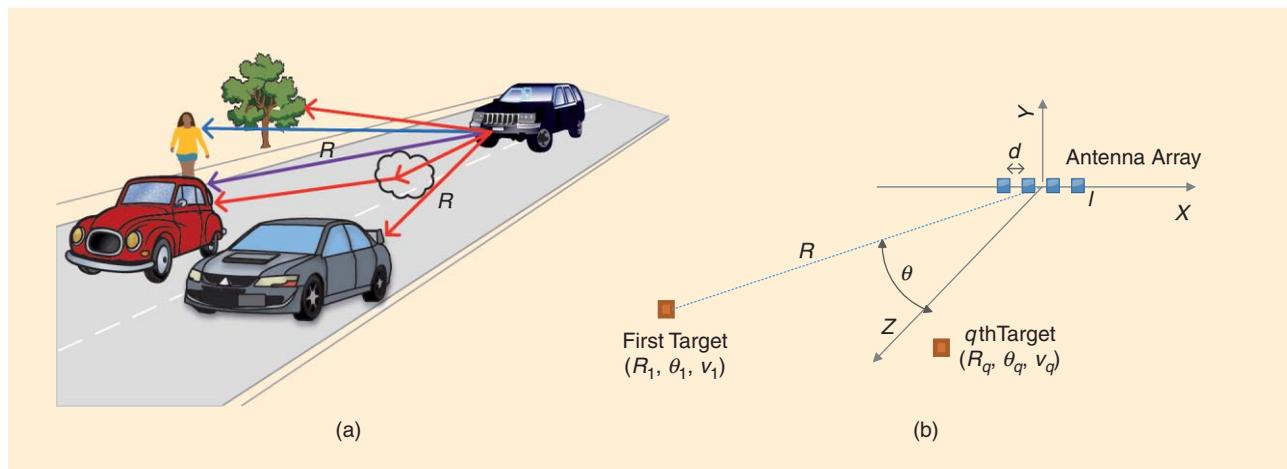


FIGURE 4. (a) A typical traffic scenario with reflections from different targets, including two cars at the same distance R . (b) The azimuth angle estimation setup using uniform linear antenna array.

unique range gate by the application of second Fourier transform across the slow time. A range-Doppler map can be found efficiently by using 2-D fast Fourier transform (FFT) (5). A demonstrative example based on the aforementioned discussion is shown in Figure 3.

Direction estimation

Use of wideband pulses such as FMCW provides discrimination of targets in both distance and velocity. The discrimination in direction can be made by means of an antenna array. Figure 4(a) depicts a realistic traffic scenario with several targets surrounding the radar that collects direct and multipath reflections from them. In such cases, to spatially resolve equidistant targets and deliver comprehensive representation of the traffic scene, angular location of targets should be estimated. Therefore, in automotive radars, the location of a target is often described in terms of a spherical coordinate system (R, θ, ϕ) , where (θ, ϕ) denote azimuthal and elevation angles, respectively. However, in this case, the single antenna radar setup as used in the range-velocity estimation problems may not be sufficient, since the measured time delay $\tau = (2(R \pm vt)/c)$ lacks the information in terms of angular locations of the targets.

To enable direction estimation, the radar should collect reflected wave data across multiple distinct dimensions. For example, locating a target using EM waves in 2-D requires the reflected wave data from the object to be collected in two distinct dimensions. These distinct dimensions can be formed in many ways using combinations of time, frequency, and space. For instance, a linear antenna array and wideband waveforms such as FMCW form two unique dimensions [13], [14]. Additionally, smaller wavelengths in mm-wave bands correspond to smaller aperture sizes and, thus, many antenna elements can be densely packed into an antenna array. Hence, the effective radiation beam, which is stronger and sharper, in turn increases the resolution of angular measurements.

Consider an antenna array located in plane $z = 0$, and let l be the abscissa corresponding to each receiver antenna position

[see Figure 4(b)]. Let (R_q, θ_q) be the position of the q th target in spherical coordinates, moving with velocity v_q relative to the radar. With the help of far field approximation [15], for the q th target, the round-trip time delay between a transmitter located at the origin and the receiver positioned at coordinate l is given by

$$\tau_{lq} = \frac{2(R_q + v_q t) + ld \sin \theta_q}{c}, \quad (6)$$

where d is the distance between antenna elements (usually half the wavelength) arranged in a linear constellation. Combining (5) and (6) gives the three-dimensional (3-D) FMCW radar output signal, which enables estimation of range, velocity, and angle. For Q number of targets, the signal can be represented as

$$d(l, n, p) \approx \sum_{q=0}^{Q-1} \alpha_q \exp \left\{ j2\pi \left[\left(\frac{2K R_q}{c} + f_{dq} \right) \frac{n}{f_s} + \frac{f_c l d \sin \theta_q}{c} + f_{dq} p T_0 + \frac{2f_c R_q}{c} \right] \right\} + \omega(l, n, p), \quad (7)$$

where α and ω correspond to same quantities as explained in the range estimation problem. The delay term τ_{lq} creates uniform phase progression across antenna elements, which permits the estimation of the angle by FFT in spatial domain, as shown in (7). Thus, 2-D location (range and angle) and speed of targets can be jointly estimated by 3-D FFT. The target location and velocity estimation problems are revisited later in the section “Advanced Estimation Techniques” with more emphasis on the high resolution algorithms and computational complexity analysis.

Radar waveforms

Various automotive radar classes, summarized in Table 1, have diverse specifications in terms of several fundamental radar system performance metrics, such as range resolution, velocity resolution, angular direction, SNR, and the probability of target detection. The type of waveform employed by a

Table 2. Radar waveforms.

Waveform Type	Transmit Waveform $s(t)$	Detection Principle	Resolution	Comments
CW	$e^{j2\pi f_c t}$	Conjugate mixing	$\Delta f_d = 1/T$	No range information
Pulsed CW	$\Pi(T_p) e^{j2\pi f_c t}$	Correlation	$\Delta R = cT_p/2 \quad \Delta f_d = 1/T_p$	Range-Doppler performance tradeoff
FMCW	$e^{j2\pi(f_c + 0.5K)t}$, $K = \frac{B}{T_0}$	Conjugate mixing	$\Delta R = c/2B \quad \Delta f_d = 1/PT_0$	Both range and Doppler information
SFCW	$e^{j2\pi f_n t}$, $f_n = f_c + (n-1)\Delta f$	Inverse Fourier transform	$\Delta R = c/2B \quad \Delta f_d = 1/PT_0$	Δf decides maximum range
OFDM	$\sum_{n=0}^{N-1} I(n) e^{j2\pi(f_c + n\Delta f)t}$	Frequency domain channel estimation	$\Delta R = c/N\Delta f \quad \Delta f_d = 1/PT_N$	Suitable for vehicular communication

B denotes bandwidth of the radar. T is the amount of time for which data is captured. N stands for a number of samples in CW and number of carriers in OFDM. $\Pi(T_p)$ is rectangular pulse of duration T_p . P is number of FM/SF-CW or OFDM blocks of duration T_0 and T_N , respectively. $I(n)$ is arbitrary sequence and Δf is carrier/frequency separation in OFDM/SFCW.

radar is a major factor that affects these metrics. The radar waveforms, as summarized in Table 2, can be characterized whether or not they are CW, pulsed and frequency, or phase modulated. Modulated radar waveforms include FM CW, stepped frequency (SF) CW, orthogonal frequency-division multiplexing (OFDM), and frequency shift keying (FSK). Each waveform type has a certain advantage in processing, implementation, and performance as follows:

- In the CW radar, a conjugate mixing of a high-frequency transmitted and received signal produces the output signal at the Doppler frequency of the target. The resolution of frequency measurement is inversely proportional to the time duration of the signal capture. The continuous nature of the waveform precludes round-trip delay measurement, which is necessary for range estimation of the target [see Figure 5(a)].

Hence, apart from ease of implementation and ability to detect target speed, the CW radar cannot provide the range information.

- Pulsed CW radar can estimate the range information as explained previously in the section “Basic Automotive Radar Estimation Problems.” The Doppler frequency can be estimated by making each pulse longer and measuring the frequency difference between the transmitted and received pulses. As shown in Figure 5(b), the pulse duration and pulse repetition frequency (PRF) are the key parameters in designing pulsed CW radar with desired range and velocity resolution.
- FMCW, also known as *linear frequency modulation (LFM)* or *chirp*, is used for simultaneous range and velocity estimation (refer to the “Velocity Estimation”

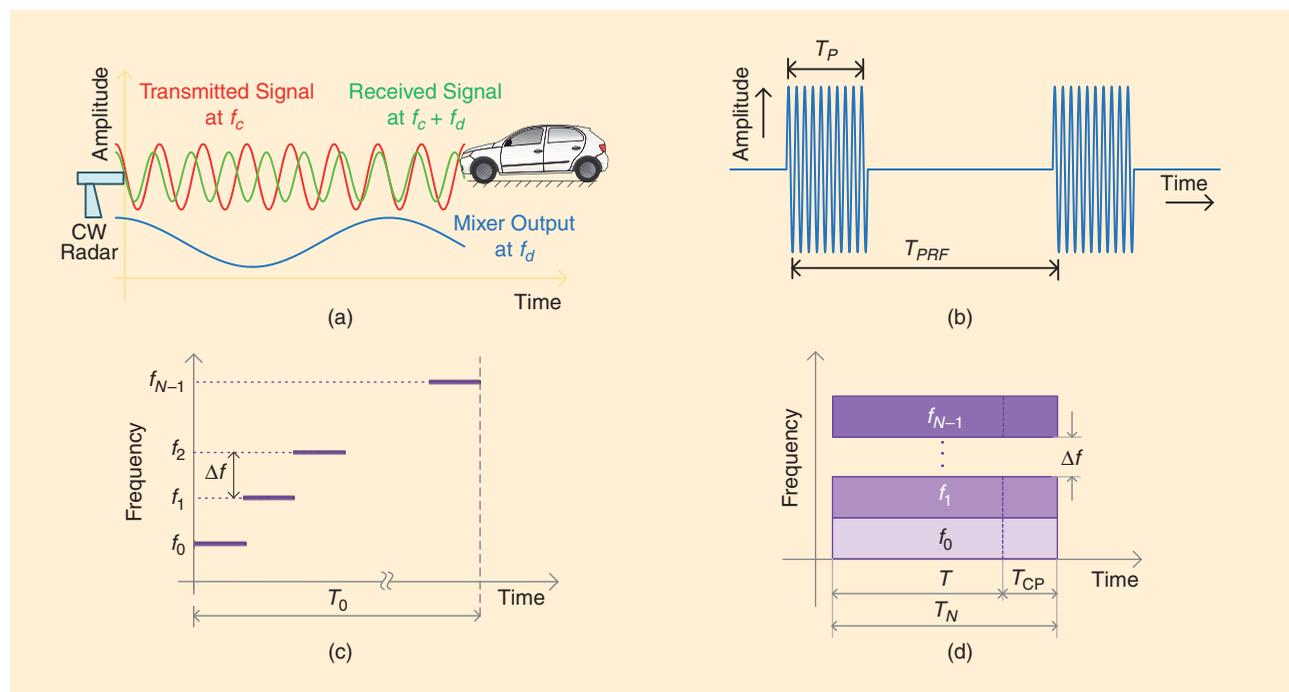


FIGURE 5. (a) Doppler frequency measurement with the CW radar. (b) A pulsed CW radar waveform with pulse repetition time T_{PRF} and pulsewidth T_p . (c) An SFCW signal with period T_0 . (d) An OFDM block with symbols time T and cyclic prefix time T_{CP} .

section for details). Due to the pulse compression, the range resolution is inversely proportional to the bandwidth of the FMCW signal and is independent of pulsewidth. For example, the short-range FMCW radar uses ultrawideband (UWB) waveforms to measure small distances with higher resolution. The Doppler resolution is a function of pulsewidth and the number of pulses used for the estimation. Thus, with the ability to measure both range and speed with high resolution, FMCW radar is widely used in the automotive industry.

- In contrast to FMCW waveforms, the frequency of FSK and SFCW varies in a discrete manner [see Figure 5(c)]. In this case, the range profile of the target and the data collected at discrete frequencies form the inverse Fourier transform relationship. Also, hybrid waveform types can be employed to achieve additive performance. FSK waveform can be combined with multislope FMCW waveform to overcome ghost targets in radar processing [16]. Similarly, alternate pulses of CW and FMCW are used to accurately estimate range and Doppler [17].
- OFDM can be viewed as another multifrequency waveform that offers unique features of the joint implementation of automotive radar and vehicle-to-vehicle communications [18], [19]. For the radar operation, the orthogonality between OFDM subcarriers is ensured by choosing carrier spacing more than maximum Doppler shift, and the cyclic prefix duration is selected greater than the longest round-trip delay [see Figure 5(d)]. The range profile is estimated through frequency domain channel estimation. OFDM radar processing along with simulation results is explained in [20].

Based on the knowledge of target statistics, radar waveforms can be optimized. Radar waveform design is revisited along with multiple-input, multiple-output (MIMO) radars in the “MIMO Radar” section.

Advanced estimation techniques

Advancements in silicon semiconductor technology have had the profound impact on the design of automotive radar systems, providing higher integration and performance at lower cost. This section reviews some sophisticated radar signal processing algorithms, which have become feasible with such advancements, especially for real-time implementation. In this section, most commonly used FMCW radar architecture is assumed and targets are considered to be stationary. Hence, (7) is reduced to a range-azimuth estimation problem with the signal model given by

$$d(l, n) \approx \sum_{q=0}^{Q-1} \alpha_q \exp \left\{ j2\pi \left[\frac{2KR_q}{c} \frac{n}{f_s} + \frac{f_c l d \sin \theta_q}{c} + \frac{2f_c R_q}{c} \right] \right\} + \omega(l, n). \quad (8)$$

The traffic imaging problem can be turned into a classical parameter estimation problem so that superresolution techniques such as MUSIC can be applied.

To elucidate advanced estimation techniques, the dimensionality of the problem is reduced to two dimensions. It should be noted that the discussed techniques can be extended to four-dimensional problems with mobile targets and elevation direction.

As discussed previously, the 2-D FFT of (8) can provide joint estimation of distance and angle. The FFT-based estimation has the

least complexity of implementation, which is $O(LN \log LN)$, where N is the number of time domain samples and L denotes the number of elements in a one-dimensional (1-D) antenna array. However, the resolution of Fourier techniques is dictated by the Rayleigh limit. While the higher range resolution can be obtained with larger FMCW bandwidth, the higher angular resolution requires more antenna elements, adding to the cost of RF front end. Additionally, the radar has to process a larger set of signal samples. However, it is important to reduce the computational load while realizing the desired angular and range resolution. We first visit the ML formulation of joint estimation of range and direction of targets. Then, we review the so-called superresolution techniques as suboptimal and lower complexity alternatives to the ML estimator.

ML estimation

The complex Gaussian observation noise in (8) is assumed to be temporally and spatially independent. ML estimation of 2-D parameters (R, θ) can be found solving the following equation:

$$\min_{R_q, \theta_q} \sum_{l=0}^{L-1} \sum_{n=0}^{N-1} \left| d(l, n) - \sum_{q=0}^{Q-1} \alpha_q \exp \left\{ j2\pi \left[\frac{2KR_q}{c} \frac{n}{f_s} + \frac{f_c l d \sin \theta_q}{c} + \frac{2f_c R_q}{c} \right] \right\} \right|^2. \quad (9)$$

Thus, depending on the granularity of (R, θ) search space, the ML estimator can offer the resolution beyond the Rayleigh limit set by system parameters such as bandwidth and number of antenna elements. However, the complexity of implementing this algorithm depends on the cardinality of the search-space as well as the number of targets. Since (R_q, θ_q) are continuous parameters, the computational complexity of ML algorithm $O(|(R, \theta)|^Q)$ becomes prohibitive. In the subsequent paragraphs, the superresolution techniques that can achieve high resolution at lower computational cost are illustrated.

Superresolution techniques

Due to their prohibitive computational cost, ML algorithms need to be implemented via suboptimal techniques. These techniques rely on collecting enough signal samples. At a sufficiently high SNR, eigenvalues and associated eigenvectors of sample covariance matrix \mathbf{C} (defined in Algorithm 1) represent the ML estimate of their true values. Hence, these eigenvectors can be used to resolve the target with high resolution.

The superresolution algorithms that rely on these techniques include multiple signal classification (MUSIC) [24] and estimation of signal parameters via rotational invariance technique (ESPRIT) [25], [26]. Recalling the 2-D (R, θ) stationary target location estimation problem, the superresolution algorithms can be applied across each dimension separately. However, this approach might lead to the so-called association problem [15]. Since the association of estimated parameters is the key step in interpreting and delivering results to the driver assist system, joint processing can be implemented. As (R, θ) domain is jointly searched for its entire range, the possibility of ghost targets is eliminated and unambiguous results are obtained [27].

As discussed previously, the temporal frequency of (8) gives the range, and spatial frequency corresponds to the angular position of the target. Hence, the traffic imaging problem can be turned into a classical parameter estimation problem so that superresolution techniques such as MUSIC can be applied. From (8), a 2-D matrix is formed, which has a Vandermonde structure across each dimension for a uniform linear antenna array. A 2-D joint superresolution was applied in the radar imaging context in [27] and later with FMCW

For proper transmitter spacing, the colocated MIMO radar can emulate a larger aperture phased array radar.

waveforms in [21], which is described in Figure 6 and Algorithm 1. The complexity of the 2-D joint superresolution algorithm lies in the cost of eigenvalue decomposition of covariance matrix $\mathbf{C}_{L_s N_s \times L_s N_s}$ and 2-D exhaustive search over the entire range of (R, θ) domain. Thus, traditional 2-D joint

superresolution algorithm has computational complexity of the order of $O(L_s N_s)^3$.

Larger size sampled covariance matrix makes 2-D joint superresolution algorithms difficult in practice. To deal with implementation issue of superresolution algorithms in real time, size of the observation space must be reduced.

Complexity reduction technique using beamspace projection

FFT-based estimation techniques have a low complexity of implementation. However, its resolution is limited by the radar bandwidth and number of antenna elements. On the other hand, the superresolution estimation resolves closer targets yet has higher computational complexity. Thus, there exists a trade-off between resolution and complexity. To reduce the computational complexity of superresolution algorithms and maintain their resolution capability, we propose two-stage estimator using a beamspace superresolution algorithm, which breaks the large problem into smaller problems using initial FFT processing [28], [29].

The computational cost of a joint superresolution algorithm lies mainly in the eigenvalue decomposition of large sample covariance matrix. Thus, to reduce the cost, the size of the covariance matrix must be reduced. Hence, as the first stage of a two-stage FFT-based-beamspace algorithm, we obtain the FFT of 2-D matrix $\mathbf{D}_{L \times N}$ in (8). From the output of this low-resolution 2-D FFT, we can determine temporal and spatial frequencies, which correspond to the approximate location of a target or cluster of targets. Once the frequencies of interest are known, we can project the data from the higher-dimensional subspace of $\mathbf{D}_{L \times N_s}$ to the lower subspace of our interest $\mathbf{D}_{L_b \times N_b}$ using DFT matrices, which form nonoverlapping beams in range and angular domain. Thus, the superresolution algorithm operates on the smaller data set, and the

Algorithm 1: The 2-D joint superresolution algorithm.

Input: Data collected using FMCW radar with stationary targets [refer to (8)] is arranged in a 2-D matrix $\mathbf{D}_{L \times N}$.

Output: a 2-D image with range and angle superresolution.

- 1) Apply spatial smoothing to remove the correlation in the reflected signal data: Vectorize each sub matrix $\mathbf{D}_{L_s \times N_s}$, which is selected using window into a column vector $\hat{\mathbf{D}}_{L_s N_s \times 1}$. For each sub matrix, find sample covariance matrix $\mathbf{C}_{L_s N_s \times L_s N_s} = (\hat{\mathbf{D}} \hat{\mathbf{D}}^H / N)$. Average the covariance matrix across possible overlapping windows (see Figure 6). This step is necessary for the application of the MUSIC algorithm, which typically assumes uncorrelated sources.
- 2) Perform the eigenvalue decomposition of the sample covariance matrix and find the noise subspace V_ω using AIC or MDL criterion to determine the number of sources [22] [23].
- 3) Obtain steering vectors in terms of the target position

$$a(R, \theta) = \text{vec} \left\{ e^{j2\pi \left[\frac{2KR}{c} \frac{n}{f_s} + \frac{f_c d \sin \theta}{c} + \frac{2f_c R}{c} \right]} \right\}.$$

- 4) Apply the MUSIC algorithm to locate the target in the 2-D space.

$$S(R, \theta) = \frac{1}{a^H(R, \theta) V_\omega V_\omega^H a(R, \theta)}.$$

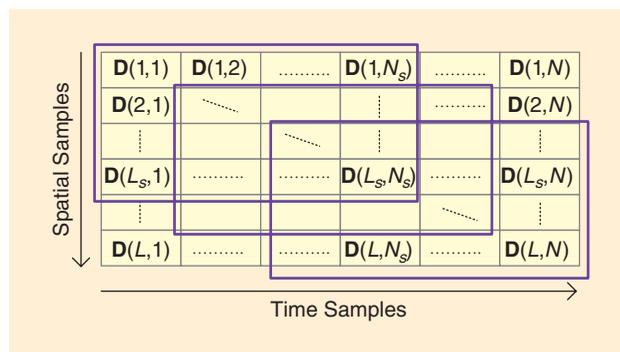


FIGURE 6. The spatial smoothing of 2-D data using a window size of $L_s \times N_s$ [21].

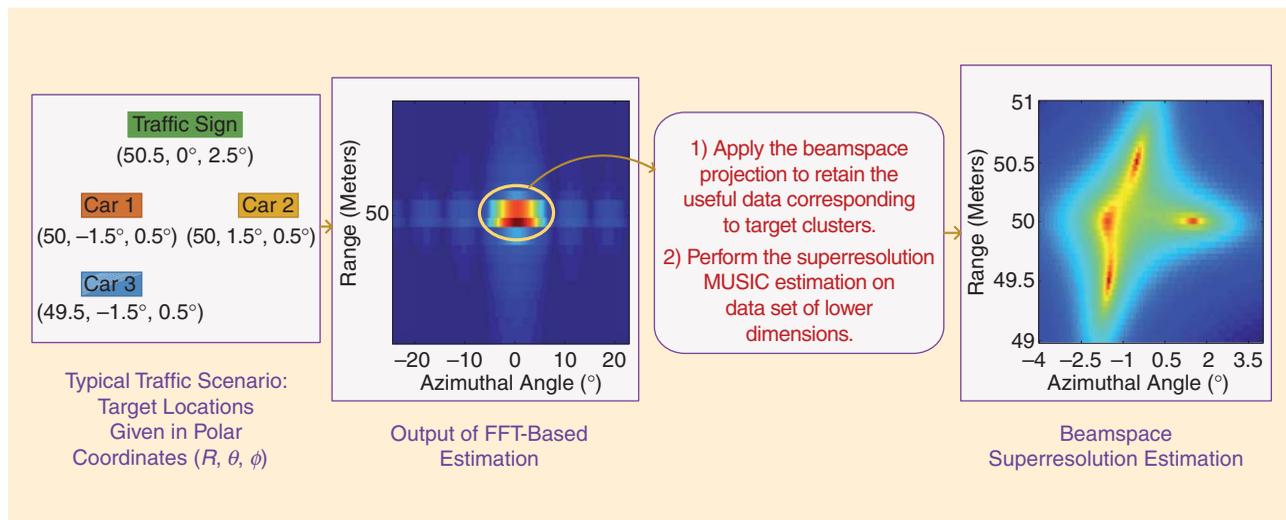


FIGURE 7. Beamspace-based superresolution estimation steps. A 77-GHz FMCW radar (SNR = 10 dB, BW = 300 MHz, $T = 300 \mu\text{s}$) is used. FFT operates on data size $[L, M] = [16, 512]$. The spatial smoothing window size is $[L_s, N_s] = [8, 50]$, and the beamspace projection reduces data to $[L_b, N_b] = [4, 5]$. Computational complexity is reduced from $\mathcal{O}(6.4 \times 10^7)$ to $\mathcal{O}(8 \times 10^3)$.

complexity of a 2-D superresolution imaging can be reduced to $\mathcal{O}(L_b N_b)^3$. Moreover, 2-D exhaustive search for the target on a finer grid operates over an area of interest, thereby further reducing the complexity. The performance of the beamspace algorithm is demonstrated in Figure 7. More detailed discussion on the complexity analysis and implementation of radar algorithms can be found in [30].

MIMO radar

MIMO radar systems employ multiple transmitters, multiple receivers, and multiple waveforms to exploit all available degrees of freedom [31]. MIMO radars can be classified as widely separated or colocated. In widely separated MIMO radar, transmit-receive antennas capture different aspects of the RCS of a target. In other words, the target appears to be spatially distributed, providing a different RCS at each antenna element. This RCS diversity can be utilized to improve the radar performance [32]. On the other hand, with colocated MIMO radar, the RCS observed by each antenna element is indistinguishable [10].

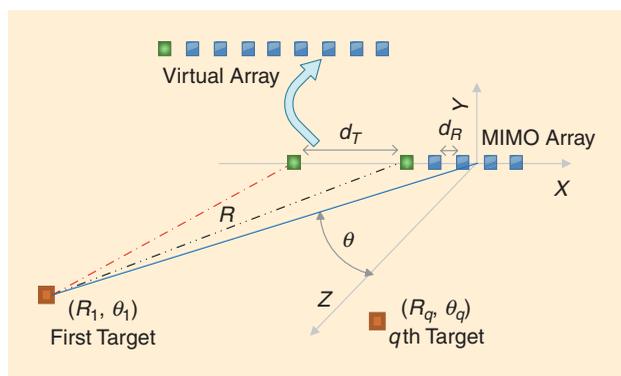


FIGURE 8. Colocated MIMO radar with a virtual array that increases the angular resolution.

Automobiles typically use colocated MIMO radars, which are compact in size [33]. For proper transmitter spacing, the colocated MIMO radar can emulate a larger aperture phased array radar (see Figure 8). This larger array is called a *virtual array*. Recall of the range-azimuth estimation problem given in (8). For the MIMO radar processing, as depicted in Figure 8, a 1-D receiver array with two transmit antennas is considered. Let L_T and L_R denote a number of transmit and receive antenna elements, respectively. Suppose that d_T and d_R represent corresponding transmit and receive antenna spacings. Also, assume that transmit and receive antenna positions in Cartesian coordinates are given by l_T and l_R . Hence, the 2-D FMCW mixer output signal across fast time and aperture is given by

$$d(l_T, l_R, n) \approx \sum_{q=0}^{Q-1} \alpha_q \exp \left\{ j2\pi \left[\frac{2K R_q n}{c} \frac{n}{f_s} + \frac{f_c \{ (l_T d_T + l_R d_R) \sin \theta_q \}}{c} + \frac{2f_c R_q}{c} \right] \right\} + \omega(l_T, l_R, n). \quad (10)$$

From (10), it is evident that if $d_T = L_R \times d_R$, then MIMO radar imitates a regular 1-D array radar with single transmit and $L_T \times L_R$ receive antenna elements. This is known as *virtual array representation*. Hence, the spatial resolution of FFT-based target imaging can be improved by the factor of L_T . With virtual array representation and substituting $l = l_T \times L_R + l_R$, the expressions similar to (8) can be obtained and the estimation algorithms discussed in the sections “Basic Automotive Radar Estimation Problems” and “Advanced Estimation Techniques” can be applied.

The challenging aspect of MIMO radar is the selection of waveforms. The waveforms can be made orthogonal in the frequency, time, or code domain [34], [35]. Consequently, the matched filter design at the receiver varies, which is necessary to separate the reflected waveforms originating from different transmitters. From the FMCW radar signal given in (4),

various orthogonal waveforms can be constructed in the following manner [36]:

- Beat frequency division: $s(t) = e^{j2\pi[(f_c - \Delta f_b)t + 0.5Kt^2 + 0.5(\Delta f_b/K)t]}$. Here, Δf_b is the frequency offset introduced for waveforms orthogonalization. The last term in the exponential corresponds to residual video phase compensation, which is necessary for coherent receiver processing.
- Modulation constant division: $s(t) = e^{j2\pi(f_c + 0.5[K + \Delta K])t}$. The modulation constant or chirp rate offset is given by ΔK , which is obtained by varying the pulse period. The bandwidth at each transmitter remains the same to maintain the range resolution. The reset time between the pulses ensures the synchronization at the receiver.
- Code division: $s(t) = e^{j[2\pi(f_c + 0.5Kt) + 0.5\beta(t)]}$, where $\beta(t)$ corresponds to the binary phase-shift keying (BPSK) signal with a low update rate that assumes values ± 1 . The bandwidth of the BPSK signal is kept smaller to ensure the proper operation of the FMCW radar.

Following the waveform selection, the waveform design can be used for further optimization of the radar performance. For the wideband radar waveforms with high-range resolution, a planar target appears to be a cluster of point targets. The extended target exhibit random reflectivity (impulse response) as its reflection consists of several waveforms added together. From the known extended target statistics, the transmitted waveform can be adapted (see Figure 9). The mutual information between a random extended target and the reflected received signal is used to optimize the radar waveform [38]. Under the constraint on the transmit power, the waveforms can be designed to minimize the mean square error in the target impulse response estimation. The solution to this problem consists of water-filling power allocation, distributing more power to target exhibiting significant scattering [39]. As shown in [40], multiuser MIMO principles can be applied to waveform design in the context of multiple target estimation and tracking.

Robust estimation techniques

So far, we have assumed that the automotive radars only receive the reflection from the targets of interest such as a vehicle traveling in front. However, in addition to direct reflections from the target of interest, the radar also receives reflections from the road debris, guard rails, and walls. This unwanted return at the radar is called *clutter*. The amount of clutter in the system changes as the surrounding environment of the vehicle varies. Hence, adaptive algorithms such as constant false alarm rate (CFAR) processing and space-time adaptive processing (STAP) can be used to mitigate the effect of clutter.

To identify valid targets in the presence of clutter, the threshold for the target detection should be properly chosen. If the amplitude of the spectrum at an estimated range is greater than some threshold, the target is said to be detected. Thus, the threshold should depend on the noise or in other words on the clutter in the given system. As clutter

To identify valid targets in the presence of clutter, the threshold for the target detection should be properly chosen.

increases, a higher threshold may be chosen. A simple CFAR method based on cell averaging can use a sliding window to derive the local clutter level by averaging the multiple range bins. As multiple targets make this detection method intricate, sophisticated techniques based on

ordered statistics can be used [41], [42].

STAP is another technique that can robustify target position estimation [43], [44]. The key idea is to use an adaptive filter that selects the target amid clutter from road and other objects. The weights of the filter change adaptively with clutter statistics. In FMCW radar (7), this filter operates on the mixer output across different chirps (i.e., P slow time samples) as well as across spatial domain (L samples from 1-D aperture). The clutter statistics are recorded with the interference covariance matrix $\mathbf{C}_{LP \times LP}$, which is calculated by averaging over the range bins surrounding the target of interest. Let $e_{LP \times 1}(\theta_t, f_{d_t})$ be the spatio-temporal steering vector pointing to the possible target. The weights of space-time adaptive filter are given by minimum variance distortionless (MVD) beamformer [44] as

$$w(\theta_t, f_{d_t}) = \frac{\mathbf{C}^{-1} e(\theta_t, f_{d_t})}{\sqrt{e^H(\theta_t, f_{d_t}) \mathbf{C}^{-1} e(\theta_t, f_{d_t})}}. \quad (11)$$

The presence of target is then tested by passing the spatio-temporal data through the filter with coefficients $w(\theta_t, f_{d_t})$. This process is conducted for all possible targets of interest.

Additionally, STAP can benefit from extra degrees of freedom in MIMO radar by using multiple transmitter antenna elements to reduce the clutter. The MIMO radar with increased virtual array size can process both direction of arrival and departure information, which shows mismatch if the signal is reflected from the clutter [11], [45].

Target tracking problem

Target tracking is an essential part of the ADAS subsystems such as collision avoidance and lane assist. In the tracking, a state (x, y, z, v_x, v_y, v_z) , which indicates the 3-D position of the

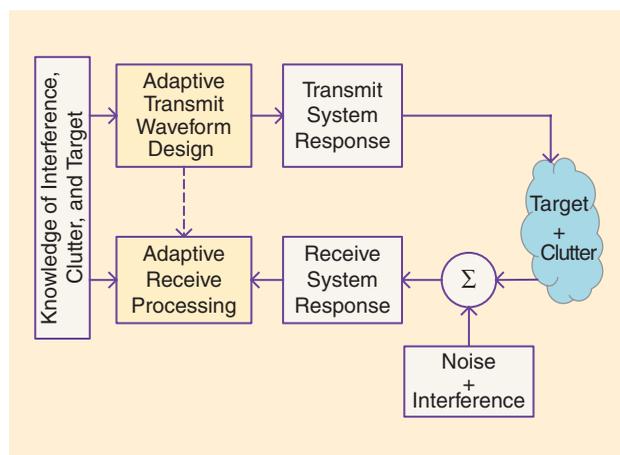


FIGURE 9. The functional block diagram of adaptive waveform design [37].

target in Cartesian coordinates and corresponding directional velocities is determined based on the current observation (R, θ, ϕ) and previous state information.

A key step in tracking is to associate separately estimated parameters of Q targets, particularly velocities (v_1, v_2, \dots, v_Q) and ranges (R_1, R_2, \dots, R_Q) with each other $[(R_1, v_1), (R_2, v_2), \dots, (R_Q, v_Q)]$. After linking estimated parameters with targets, the targets are associated with tracks. For example, if each target follows a separate track, then there are Q tracks in the system. The association problem becomes complex when two tracks cross each other. Different methods to perform data association include joint probabilistic data association (JPDA), nearest neighbor (NN), and fuzzy logic [46].

Following the data association, tracking can be performed using well-known algorithms such as Kalman filtering. For each track, a separate filter is implemented. These filters operate in parallel. Since the observation vector (R, θ, ϕ) has a nonlinear relationship with the state vector (x, y, z, v_x, v_y, v_z) , an extended Kalman filter (EKF) is used. The state equation that captures the effect state transition over time [47] is given by

$$\underbrace{\begin{bmatrix} x[n] \\ y[n] \\ v_x[n] \\ v_y[n] \end{bmatrix}}_{s[n]} = \underbrace{\begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_A \underbrace{\begin{bmatrix} x[n-1] \\ y[n-1] \\ v_x[n-1] \\ v_y[n-1] \end{bmatrix}}_{s[n-1]} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ u_x[n] \\ u_y[n] \end{bmatrix}}_{\text{State Perturbation}}, \quad (12)$$

where T is the observation interval. The observation vector is related with state vector via

$$\underbrace{\begin{bmatrix} R[n] \\ \theta[n] \end{bmatrix}}_{x[n]} = \underbrace{\begin{bmatrix} \sqrt{x^2[n] + y^2[n]} \\ \arctan \frac{y[n]}{x[n]} \end{bmatrix}}_{h(s[n])} + \underbrace{\begin{bmatrix} \omega_R[n] \\ \omega_\theta[n] \end{bmatrix}}_{\text{Observation noise}}. \quad (13)$$

From the knowledge of the previous state, the present state is predicted based on the state equation (12). Using (13) and the present observation, the predicted value is updated. The amount of correction depends on the SNR of the observations; see [47] for more details. Vehicle tracking problems are also addressed in [48]–[51].

Pedestrian detection

Pedestrian, bike, and wild life detection is essential for a driver assist and collision avoidance system. As a pedestrian walks, a small change in range produces very low Doppler shift. In other words, the micromotion of a target produces what is known as a *micro-Doppler* [52]. Likewise, the periodic motion of limbs creates a periodic pattern in velocity over time, which is also known as the *micro-Doppler signature*. This signature, along with other feature extraction and matching algorithms, can be used to uniquely identify pedestrian walking. More details about an analysis of human gait using range-Doppler plots are given in [53].

A realistic simulation setup should include radiation patterns of the transmit and receive antenna elements, which count for the direction dependent scaling of the transmitted and reflected signals according to the geometry of the system.

Moreover, the pedestrian detection task becomes more challenging due to a smaller RCS of the human body [54]. To make the pedestrian detection robust, the radar based on micro-Doppler estimation can be combined with inputs from a vision sensor [55]. Also, the tracking algorithms discussed previously can help predict pedestrian movement [56].

Let us discuss how the micro-Doppler signature is extracted using FMCW radar processing in (5). First, 2-D signal samples obtained across slow and fast time are converted into single dimensional signals by

range gating. Typically, FFT is performed across fast time n and only the frequency corresponding to the range of interest R_0 is retained (assume single target with micromotion at R_0). Neglecting range-Doppler coupling and effect of finite length FFT, (5) can be rewritten as

$$\hat{d}(p) \approx \alpha_0 \exp \left\{ j2\pi \left(\frac{2f_c}{c} [R_0 + \Omega(pT_0)] \right) \right\} + \hat{w}(p), \quad (14)$$

where $\Omega(\cdot)$ is the function characterizing the micromotion of the target. As explained in [52], the short-time Fourier transform (STFT) of (14) gives the instantaneous variation of Doppler across time. Detail analysis regarding micro-Doppler vibration measurements using FMCW radar is done in [57]. In addition to pedestrian detection, micro-Doppler also can be used to identify the type of a vehicle (truck, sedan, etc.) by characterizing its vibration pattern on top of Doppler shift produced by its bulk motion [58], [59].

FMCW radar EM simulation setup

Radar algorithms are often verified by means of simulations, which reduces the cost of prototyping and testing. While modeling the radar systems, the targets and channels under consideration are assumed to be ideal. The targets are modeled as objects with perfect reflectivity, and the signals are assumed to propagate through unobstructed paths. To verify the viability of various radar estimation algorithms in the real world, it is necessary to use computational EM software to simulate potential target RCSs and channels.

A realistic simulation setup should include radiation patterns of the transmit and receive antenna elements, which count for the direction dependent scaling of the transmitted and reflected signals according to the geometry of the system. In addition, EM waves undergo reflection, diffraction, and scattering, depending on the shape and size of the target with respect to its wavelength. To incorporate these phenomena, Maxwell's equations with appropriate boundary conditions must be solved. Along with numerical computing, software packages such as MATLAB or MATHEMATICA and EM simulators such as ADS [60], FEKO [61] or Xpatch [62] can be used for the accurate modeling of the automotive radar imaging. The effect of RF impairments such as phase noise, local oscillator leakage, and in-phase and quadrature imbalance can be modeled either in MATLAB or an EM simulator such as ADS.

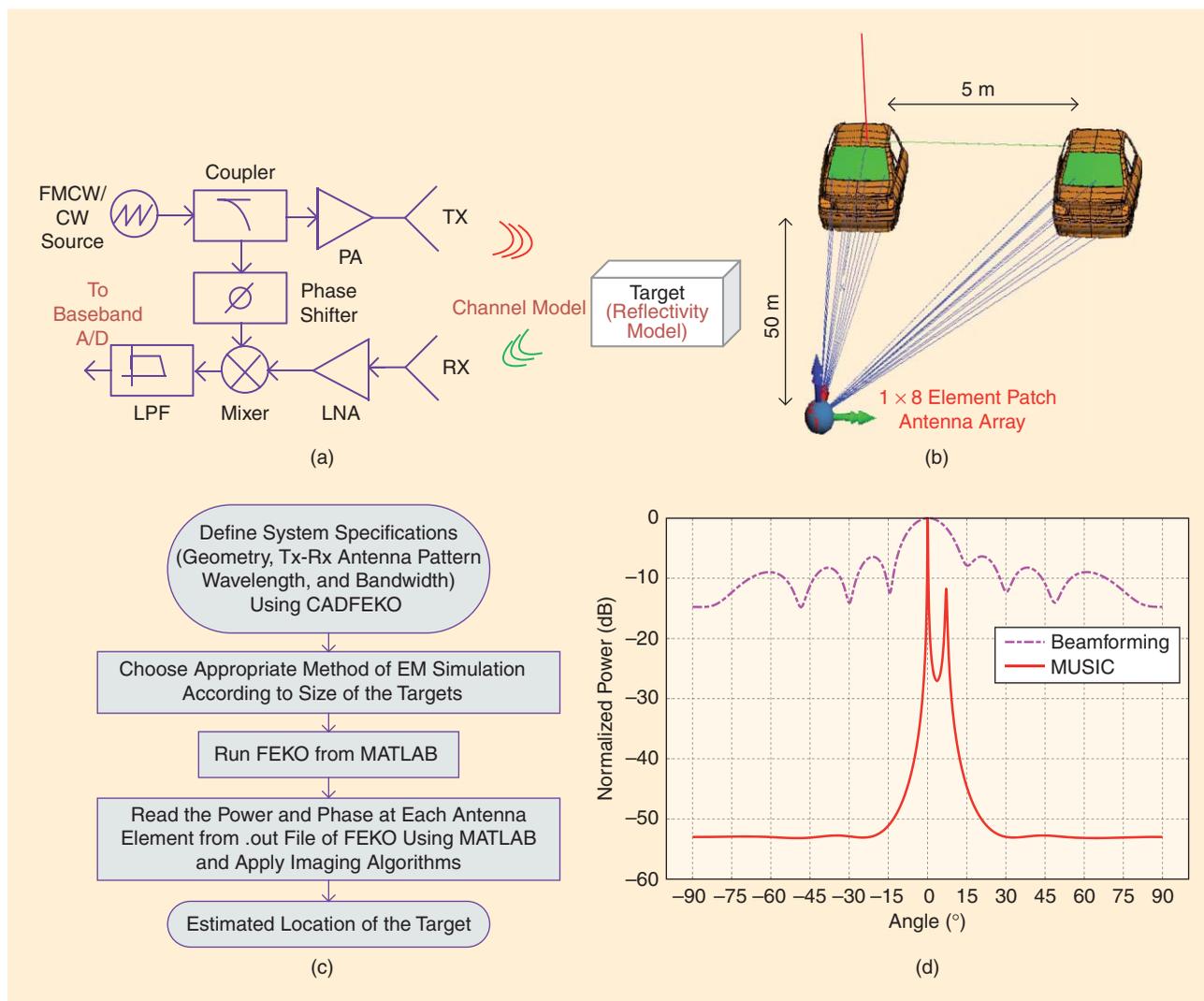


FIGURE 10. (a) An RF block diagram showing various aspects of realistic radar simulations. (b) The antenna array, channel, and target modeling using ray tracing in FEKO. (c) The algorithm for FEKO and MATLAB integration. (d) Azimuthal angle estimation for the scenario in Figure 10(b) using MUSIC and beamforming algorithms; it shows the effect of directional RCS of the car.

We demonstrate a realistic automotive radar simulation setup based on FEKO and MATLAB implementations, as illustrated in Figure 10.

Data fusion and challenges

The automotive radar output is often combined with outputs from other sensors such as lidar, camera vision, and ultrasound. Lidar and vision sensors can help enhance discrimination capabilities and reduce computation costs by delivering faster response. Independent observations from other sensors must be combined with radar systems to increase the reliability. For example, the lidar provides improved target detection on curved roads. Radar offers superior speed measurements, as they rely on the Doppler effect as opposed to lux measurement in lidar [63]. Moreover, lidar is more sensitive to environmental factors such as snow, fog, dust, and rain [64].

When multiple sensors are in operation, all measurements should be synchronized to a common clock using time stamp-

ing. Observations from individual sensors are typically combined together to form global sensor data. The relative placement, orientation, and mathematical models of each sensor should be considered. Details about fusion techniques such as object-list-level, track-to-track, low-level, and feature-level fusion are discussed in [65] and [66]. More information about real-time object detection using learning algorithms can be found in [67].

Another important aspect of automotive radars is the interference between two vehicles [68]. Analytical studies point out reduced radar sensitivity in such cases. Null steering, tracking, coded sequences, and interleaving are among several techniques used for interference mitigation. An additional feature of the intelligent transportation system can include vehicle-to-vehicle communication, which can also help to avoid collision [69], [70].

Conclusions

As we progress toward fully autonomous driving, many challenges and innovative solutions will emerge. The fundamental

component of these autonomous systems is the automotive radar, which has become feasible due to prospering mm-wave circuit technology. Concurrently, sophisticated signal processing techniques have gained momentum to efficiently utilize the automotive radar hardware. In this article, we have presented various signal processing aspects of automotive radars, starting from basics of range and velocity estimation to complex 3-D end-to-end EM simulation. The target location estimation techniques are explained with sufficient mathematical details and illustrative examples so that the article may also serve as a tutorial. For briefly touched-on advanced topics in the field, we have pointed to relevant literature, which readers can pursue according to their interests. This review article should help researchers and engineers take a first step forward in developing novel automotive radar signal processing techniques.

Acknowledgment

This work was, in part, supported by Semiconductor Research Corporation through the Texas Analog Center of Excellence at the University of Texas at Dallas under task 1836.150.

Authors

Sujeet Patole (sujeet.patole@utdallas.edu) received his B.E. degree in electronics and telecommunication engineering from the University of Pune, India, in 2009 and his M.S. degree in electrical engineering from the University of Texas at Dallas in 2013. He is a recipient of the Jonsson School Graduate Scholarship and Certificate of Academic Achievement Award from the University of Texas at Dallas, where he is currently pursuing his Ph.D. degree in electrical engineering. During the summers of 2014 and 2015, he worked at Texas Instruments, developing state-of-the-art signal processing algorithms. His research interests include millimeter-wave imaging, radar signal processing, and multiple antenna communication systems. He is a Student Member of the IEEE.

Murat Torlak (torlak@utdallas.edu) is a professor in the Department of Electrical and Computer Engineering at the University of Texas at Dallas. He received his B.S. degree from Hacettepe University, Turkey, in 1992 and his M.S. and Ph.D. degrees in electrical engineering from the University of Texas at Austin in 1995 and 1999, respectively. He was a visiting scholar at the University of California Berkeley in 2008. His current research focus is on experimental platforms for multiple antenna systems, millimeter-wave systems, signal processing techniques for automotive radars, adaptation in wireless communication systems, and signal processing techniques of radio-frequency/analog imperfections in communications transceivers. He was an associate editor of *IEEE Transactions on Wireless Communications* during 2008–2013. He was the general chair of the Symposium on Millimeter Wave Imaging and Communications of the 2013 IEEE Global Conference on Signal and Information Processing. He is a Senior Member of the IEEE.

Dan Wang (danwang1981@ti.com) received her B.S. degree in radio engineering from Southeast University, China,

in 2006 and her M.S. degree in electrical engineering from the University of Minnesota–Twin Cities in 2009. She joined Texas Instruments in 2012 and is now a senior system engineer in the field of radar signal processing and system design. She has been developing radar applications in the automotive and various industry areas.

Murtaza Ali (murtaza.ali@ieee.org) received his B.S. degree in electrical engineering from Bangladesh University of Engineering and Technology in 1989 and his Ph.D. degree in electrical engineering from the University of Minnesota–Twin Cities in 1995. He is currently director of systems engineering at Uhnder, Inc. Prior to joining Uhnder, he was a distinguished member of technical staff and the manager of the Perception and Analytics Lab at Texas Instruments. His current research interests include environment sensing and perception using technologies such as radar, lidar, and cameras. He holds 18 U.S. patents and has published more than 40 papers in refereed and invited forums. He is a Senior Member of the IEEE.

References

- [1] M. Richards, M. A. Richards, J. A. Scheer, and W. A. Holm, *Principles of Modern Radar: Basic Principles*. Institution of Engineering and Technology, 2010.
- [2] W. D. Jones, "Keeping cars from crashing," *IEEE Spectr.*, vol. 38, no. 9, pp. 40–45, Sept. 2001.
- [3] E. Guizzo, "How Google's self-driving car works," *IEEE Spectr. Online*, vol. 18, Oct. 2011.
- [4] H. H. Meinel, "Evolving automotive radar: From the very beginnings into the future," in *Proc. European Conf. Antennas and Propagation*, Hague, The Netherlands, 2014, pp. 3107–3114.
- [5] S. Kobashi, et al., "3D-Scan radar for automotive application," in *Proc. ITS World Congr.*, Vienna, Austria, 2012, vol. 38, pp. 3–7.
- [6] I. Gresham, N. Jain, T. Budka, A. Alexanian, N. Kinayman, B. Ziegner, S. Brown, and P. Staecker, "A compact manufacturable 76-77-GHz radar module for commercial ACC applications," *IEEE Trans. Microwave Theory Tech.*, vol. 49, no. 1, pp. 44–58, Jan. 2001.
- [7] D. M. Grimes and T. O. Jones, "Automotive radar: A brief review," *Proc. IEEE*, vol. 62, no. 6, pp. 804–822, June 1974.
- [8] J. Wenger, "Automotive radar: Status and perspectives," in *Proc. IEEE Compound Semiconductor Integrated Circuit Symp.*, Palm Springs, CA, 2005, pp. 21–25.
- [9] F. Fölster and H. Rohling, "Signal processing structure for automotive radar," *Frequenz*, vol. 60, no. 1/2, pp. 1–4, 2006.
- [10] J. Li and P. Stoica, "MIMO radar with colocated antennas," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 106–114, Sept. 2007.
- [11] J. Yu and J. Krolik, "MIMO multipath clutter mitigation for GMTI automotive radar in urban environments," in *Proc. IET Int. Conf. Radar System*, Glasgow, U.K., 2012, pp. 24–24.
- [12] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 GHz frequency band," *IEEE Trans. Microwave Theory Tech.*, vol. 60, no. 3, pp. 845–860, Mar. 2012.
- [13] M. Soumekh, "Array imaging with beam-steered data," *IEEE Trans. Image Process.*, vol. 1, no. 3, pp. 379–390, July 1992.
- [14] S. Patole and M. Torlak, "Two dimensional array imaging with beam steered data," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5181–5189, Dec. 2013.
- [15] H. Trees, *Detection, Estimation, and Modulation Theory, Optimum Array Processing*. New York: Wiley, 2001.
- [16] H. Rohling and M. M. Meinecke, "Waveform design principles for automotive radar systems," in *Proc. CIE Int. Conf. Radar Systems*, Beijing, China, 2001, pp. 1–4.
- [17] G. Wang, C. Gu, T. Inoue, and C. Li, "Hybrid FMCW-interferometry radar system in the 5.8 GHz ISM band for indoor precise position and motion detection," in *Proc. IEEE Int. Microwave Symp.*, Seattle, WA, 2013, pp. 1–4.

- [18] B. Donnet and I. Longstaff, "Combining MIMO radar with OFDM communications," in *Proc. IEEE European Radar Conf.*, Manchester, U.K., 2006, pp. 37–40.
- [19] C. Sturm, T. Zwick, and W. Wiesbeck, "An OFDM system concept for joint radar and communications operations," in *Proc. IEEE Vehicular Technology Conf.*, Barcelona, Spain, 2009, pp. 1–5.
- [20] C. Sturm, E. Pancera, T. Zwick, and W. Wiesbeck, "A novel approach to OFDM radar processing," in *Proc. IEEE Radar Conf.*, Pasadena, CA, 2009, pp. 1–4.
- [21] F. Belfiori, W. van Rossum, and P. Hoogetboom, "Application of 2D MUSIC algorithm to range-azimuth FMCW radar data," in *Proc. European Radar Conf.*, Amsterdam, 2012, pp. 242–245.
- [22] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Control*, vol. 19, no. 6, pp. 716–723, Dec. 1974.
- [23] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
- [24] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antenna Propagat.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
- [25] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 7, pp. 984–995, July 1989.
- [26] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, July 1996.
- [27] J. Odendaal, E. Barnard, and C. Pistorius, "Two-dimensional superresolution radar imaging using the MUSIC algorithm," *IEEE Trans. Antennas Propagat.*, vol. 42, no. 10, pp. 1386–1391, Oct. 1994.
- [28] M. D. Zoltowski, G. M. Kautz, and S. D. Silverstein, "Beamspace Root-MUSIC," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 344–364, Jan. 1993.
- [29] S. Patole and M. Torlak, "Fast 3D joint superresolution imaging," submitted for publication.
- [30] D. Wang and M. Ali, "Synthetic aperture radar on low power multi-core digital signal processor," in *Proc. IEEE Conf. High Performance Extreme Computing*, Waltham, MA, 2012, pp. 1–6.
- [31] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini, and R. Valenzuela, "MIMO radar: An idea whose time has come," in *Proc. IEEE Radar Conf.*, Philadelphia, PA, 2004, pp. 71–78.
- [32] A. M. Haimovich, R. S. Blum, and L. J. Cimini, "MIMO radar with widely separated antennas," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 116–129, Dec. 2008.
- [33] S. Lutz, K. Baur, and T. Walter, "77 GHz lens-based multistatic MIMO radar with collocated antennas for automotive applications," in *IEEE Microwave Symp. Dig.*, Montreal, Canada, 2012, pp. 1–3.
- [34] F. Gini, A. De Maio, and L. Patton, *Waveform Design and Diversity for Advanced Radar Systems*. London: Institution of Engineering and Technology, 2012.
- [35] A. Zwanetski and H. Rohling, "Continuous wave MIMO radar based on time division multiplexing," in *Proc. Int. Radar Symp.*, Warsaw, Poland, 2012, pp. 119–121.
- [36] J. D. Wit, W. Van Rossum, and A. D. Jong, "Orthogonal waveforms for FMCW MIMO radar," in *Proc. IEEE Radar Conf.*, Kansas City, MO, 2011, pp. 686–691.
- [37] L. K. Patton and B. D. Rigling, "Modulus constraints in adaptive radar waveform design," in *Proc. IEEE Radar Conf.*, 2008, pp. 1–6.
- [38] M. R. Bell, "Information theory and radar waveform design," *IEEE Trans. Inform. Theory*, vol. 39, no. 5, pp. 1578–1597, Sept. 1993.
- [39] Y. Yang and R. S. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 330–343, Jan. 2007.
- [40] A. Leshem and A. Nehorai, "Information theoretic radar waveform design for multiple targets," in *Proc. Int. Waveform Diversity and Design Conf.*, Pisa, Italy, 2007, pp. 362–366.
- [41] D. Kok and J. S. Fu, "Signal processing for automotive radar," in *Proc. IEEE Int. Radar Conf.*, Arlington, VA, 2005, pp. 842–846.
- [42] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 19, no. 4, pp. 608–621, July 1983.
- [43] J. Ward, "Space-time adaptive processing for airborne radar," in *Proc. Int. Conf. Acoustics, Speech and Signal Processing*, London, U.K., 1995, vol. 5, pp. 2809–2812.
- [44] M. C. Wicks, M. Rangaswamy, R. Adve, and T. B. Hale, "Space-time adaptive processing: A knowledge-based perspective for airborne radar," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 51–65, Jan. 2006.
- [45] V. F. Mecca, D. Ramakrishnan, and J. L. Krolik, "MIMO radar space-time adaptive processing for multipath clutter mitigation," in *Proc. IEEE Workshop Sensor Array and Multichannel Processing*, Waltham, MA, 2006, pp. 249–253.
- [46] A. Macaveiu, A. Campeanu, and I. Nafornita, "Kalman-based tracker for multiple radar targets," in *Proc. Int. Conf. Communication*, Bucharest, Romania, 2014, pp. 1–4.
- [47] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.
- [48] M. Z. Ikram and M. Ali, "3-D object tracking in millimeter-wave radar for advanced driver assistance systems," in *Proc. IEEE Global Conf. Signal and Information Processing*, Austin, TX, 2013, pp. 723–726.
- [49] A. Macaveiu and A. Câmpeanu, "Automotive radar target tracking by Kalman filtering," in *Proc. Int. Conf. Telecommunication Modern Satellite, Cable and Broadcasting Services*, Nis, Serbia, 2013, vol. 2, pp. 553–556.
- [50] D. Oprisan and H. Rohling, "Tracking systems for automotive radar networks," in *Proc. RADAR*, Edinburgh, U.K., 2002, pp. 339–343.
- [51] N. Floudas, A. Polychronopoulos, and A. Amditis, "A survey of filtering techniques for vehicle tracking by radar equipped automotive platforms," in *Proc. IEEE Int. Conf. Information Fusion*, Philadelphia, PA, 2005, vol. 2, pp. 1436–1443.
- [52] V. C. Chen, F. Li, S. S. Ho and H. Wechsler, "Micro-Doppler effect in radar: Phenomenon, model, and simulation study," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 1, pp. 2–21, Jan. 2006.
- [53] H. Rohling, S. Heuel, and H. Ritter, "Pedestrian detection procedure integrated into an 24 GHz automotive radar," in *Proc. IEEE Radar Conf.*, Washington, DC, 2010, pp. 1229–1232.
- [54] N. Yamada, Y. Tanaka, and K. Nishikawa, "Radar cross section for pedestrian in 76 GHz band," in *Proc. IEEE European Microwave Conf.*, Paris, France, 2005, vol. 2, pp. 4–8.
- [55] D. M. Gavrilu, "Sensor-based pedestrian protection," *IEEE Intell. Syst.*, vol. 16, no. 6, pp. 77–81, 2001.
- [56] M. Bertozzi, A. Broggi, A. Fascioli, A. Tibaldi, R. Chapuis, and F. Chausse, "Pedestrian localization and tracking system with Kalman filtering," in *Proc. IEEE Intelligent Vehicles Symp.*, Detroit, MI, 2004, pp. 584–589.
- [57] L. Ding, M. Ali, S. Patole, and A. Dabak, "Vibration parameter estimation using FMCW radar," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Beijing, China, 2016, pp. 2224–2228.
- [58] T. Thayaparan, S. Abrol, and E. Riseborough, "Micro-Doppler radar signatures for intelligent target recognition," DTIC Document, Tech. Rep., 2004.
- [59] Y. Li, L. Du, and H. Liu, "Moving vehicle classification based on micro-Doppler signature," in *Proc. IEEE Int. Conf. Signal Processing, Communications and Computing*, Xian, China, 2011, pp. 1–4.
- [60] Advanced design system. [Online]. Available: <http://www.keysight.com/en/pc-1297113/advanced-design-system-ads>
- [61] Feko Homepage [Online]. Available: <https://www.feko.info/>
- [62] D. Andersh, J. Moore, S. Kosanovich, D. Kapp, R. Bhalla, R. Kipp, T. Courtney, A. Nolan, F. German, J. Cook, and J. Hughes, "Xpatch 4: the next generation in high frequency electromagnetic modeling and simulation software," in *Proc. IEEE Radar Conf.*, Alexandria, VA, 2000, pp. 844–849.
- [63] D. Göhring, M. Wang, M. Schnürmacher, and T. Ganjineh, "Radar/lidar sensor fusion for car-following on highways," in *Proc. Int. Conf. Automation, Robotics and Applications*, Wellington, New Zealand, 2011, pp. 407–412.
- [64] R. H. Rasshofer and K. Gresser, "Automotive radar and lidar systems for next generation driver assistance functions," *Adv. Radio Sci.*, vol. 3, pp. 205–209, May 2005.
- [65] M. Mahlich, R. Hering, W. Ritter, and K. Dietmayer, "Heterogeneous fusion of video, LIDAR and ESP data for automotive ACC vehicle tracking," in *Proc. IEEE Int. Conf. Multisensor Fusion and Integration for Intelligent Systems*, Heidelberg, Germany, 2006, pp. 139–144.
- [66] B. Steux, C. Laurgeau, L. Salesse, and D. Wautier, "Fade: A vehicle detection and tracking system featuring monocular color vision and radar data fusion," in *Proc. IEEE Intelligent Vehicle Symp.*, Versailles, France, 2002, vol. 2, pp. 632–639.
- [67] P. Viola and M. Jones, "Robust real-time object detection," *Int. J. Comput. Vis.*, vol. 4, Feb. 2001.
- [68] G. M. Brooker, "Mutual interference of millimeter-wave radar systems," *IEEE Trans. Electromagnet. Compat.*, vol. 49, no. 1, pp. 170–181, Feb. 2007.
- [69] P. Papadimitratos, A. D. La Fortelle, K. Evenness, R. Brignolo, and S. Cosenza, "Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 84–95, Nov. 2009.
- [70] F. Ye, M. Adams, and S. Roy, "V2V wireless communication protocol for rear-end collision avoidance on highways," in *Proc. IEEE Int. Conf. Communications Workshop*, Beijing, China, 2008, pp. 375–379.

Florian Engels, Philipp Heidenreich, Abdelhak M. Zoubir,
Friedrich K. Jondral, and Markus Wintermantel

Advances in Automotive Radar

A framework on computationally efficient high-resolution frequency estimation

Radar technology is used for many applications of advanced driver assistance systems (ADASs) and is considered as one of the key technologies for highly automated driving (HAD). An overview of conventional automotive radar processing is presented and critical use cases are pointed out in which conventional processing is bound to fail due to limited frequency resolution. Consequently, a flexible frame-

work for computationally efficient high-resolution frequency estimation is presented. This framework is based on decoupled frequency estimation in the Fourier domain, where high-resolution processing can be applied to either the range, relative velocity, or angular dimension.

Real data obtained from series-production automotive radar sensor are presented to show the effectiveness of the presented approach.

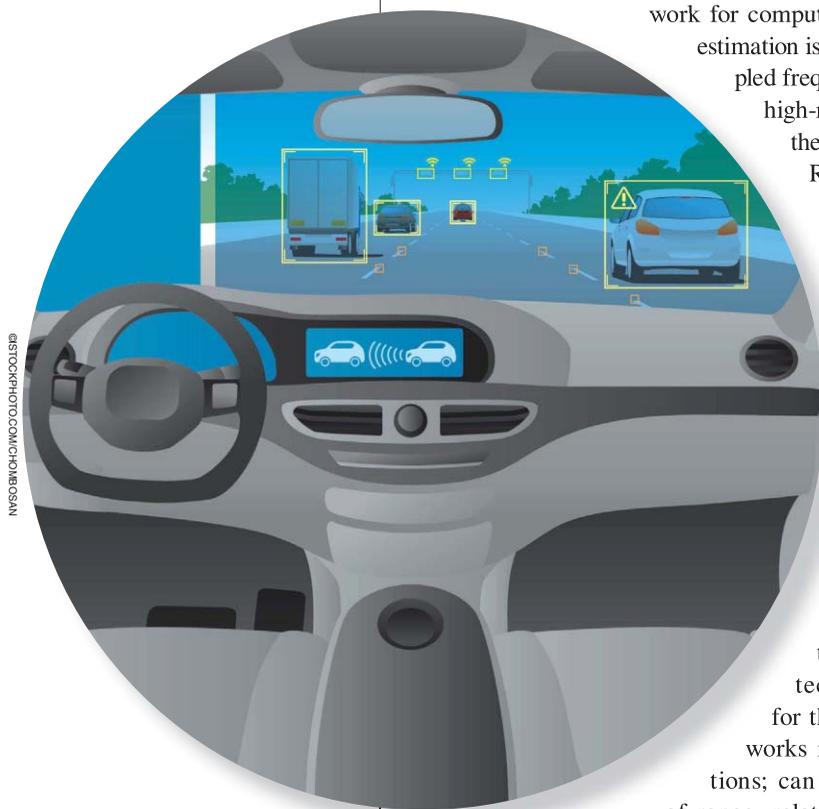
Introduction

An increasing amount of advanced signal processing algorithms is used in various automotive applications [1], [2]. Currently, one of the most dynamic topics in the automotive industry is the development of ADASs toward HAD. The performance and reliability of these systems strongly depends on the capabilities of the environmental sensing.

Radar technology has some unique advantages when compared to camera or lidar technologies and has become indispensable for the development of ADASs and HAD. Radar works reliably in bad weather and lighting conditions; can provide accurate and direct measurements of range, relative velocity, and angle of multiple targets; and can provide a high range coverage of more than 200 m.

Radar is typically used in current ADASs, such as adaptive cruise control (ACC) [3, Ch. 24], forward collision avoidance (FCA) [3, Ch. 25], lane-change assist [3, Ch. 28], or evasion assist [3, Ch. 29], to name a few. Moreover, radar is considered as one of the key technologies for HAD [4]–[6].

Classical radar preprocessing consists of pulse compression, Doppler processing, and digital beamforming. The use of chirp sequence modulation with stretch



©ISTOCKPHOTO.COM/CHIRBOSSAN

Digital Object Identifier 10.1109/MSP.2016.2637700
Date of publication: 3 March 2017

processing in state-of-the-art automotive radar sensors requires a three-dimensional (3-D) Fourier transform as a preprocessing step. The preprocessing is followed by power detection, target parameter estimation, target tracking and clustering, and, optionally, target classification, road estimation, and occupancy grid map methods, to mention a few. The main focus of this article is target parameter estimation, particularly with high-resolution capability. The aforementioned remaining steps are considered extensively in the literature and will not be discussed further in this article. An overview of target detection techniques can be found in [7]. Target tracking and clustering for automotive radars is considered in [8]–[10], target classification in [11] and [12], road estimation in [13], and occupancy grid map methods in [14].

Current limitations with respect to the requirements of the automotive industry [4] in critical-use cases include limited range and angular resolution. The range resolution is typically limited by the bandwidth, whereas the angular resolution is limited by the array aperture and the number of receive channels. Increasing bandwidth, array aperture, and the number of receive channels strongly affect hardware cost. Because the higher range and angular resolution is not required all the time, but only in critical use cases, a cost-effective alternative is to use high-resolution frequency estimation for selected processing cells.

High-resolution frequency estimation is a relatively mature field of research in signal processing theory and applications. This holds for multidimensional frequency estimation, as considered, e.g., in [15]–[17], as well as one-dimensional (1-D) frequency estimation, with the special case of high-resolution direction of arrival (DOA) estimation [18], [19]. Methods with high-resolution capability can be coarsely divided into subspace-based methods, maximum likelihood methods, and methods based on the theory of compressed sensing [20], [21]. Whereas subspace-based and maximum likelihood methods are parametric approaches, which require the number of sources to be known a priori, compressed sensing approaches solve a general nonparametric spectral estimation problem, which is independent of the number of sources. Subspace-based methods are computationally more attractive in general, whereas maximum likelihood and compressed sensing methods can be applied with a single snapshot. Note that many compressed sensing algorithms and efficient solvers have been developed recently: a study of compressed sensing DOA estimation with a single snapshot is considered in [21], recent theoretical results and an off-grid method are presented in [20]. Although we restrict ourselves to uniform sampling, we remark that when nonuniform sampling is an option and multiple snapshots are available, recent results on difference set sampling, e.g., [22], can be applied.

High-resolution DOA estimation has been considered in automotive radar [23]–[25] to meet the demands in critical use cases. Often, only a single snapshot is available for estimation so that maximum likelihood methods, or nonlinear least squares (NLS) methods, are generally favored over subspace-based methods. The main challenge in automotive

radar applications of high-resolution frequency estimation is to design algorithms that can achieve, at the same time, a high performance, can run in real time, and are memory efficient. Toward this end, we present a flexible framework and practical aspects of frequency estimation for state-of-the-art automotive radar sensors, which is computationally simple, memory efficient, and has high-resolution capability in all dimensions.

Conventional radar processing

Radar data model

State-of-the-art radar sensors use the chirp sequence modulation principle and an array of receive antennas to independently measure the range, the relative velocity, and the angle of multiple targets in the field of view [26]–[28]. After down-mixing, filtering, and sampling, the signal model for the radar measurements of a coherent processing interval (CPI) is a superposition of K 3-D complex sinusoids [29], [30]

$$x(l_s, m_s, n_s) = \sum_{k=1}^K a_k e^{j(\lambda_k l_s + \mu_k m_s + \nu_k n_s)} + \xi(l_s, m_s, n_s) \quad (1)$$

for $l_s = 0, \dots, L_s - 1$, $m_s = 0, \dots, M_s - 1$, and $n_s = 0, \dots, N_s - 1$, where K is the number of targets; a_k is the complex response parameter; λ_k , μ_k , and ν_k are the normalized radian frequencies corresponding to the range dimension, relative velocity dimension, and angular dimension, respectively; ξ is white, circular complex Gaussian noise; and L_s , M_s , and N_s are the number of range samples, pulses, and antennas, respectively. Frequency parameters λ and μ are directly proportional to the range and relative velocity, respectively, whereas for a uniform linear array (ULA), ν is related via $\kappa d \sin(\phi)$ to the azimuth angle ϕ , where κ is the wavenumber, and d is the antenna spacing.

The periodogram

An optimal approach for frequency parameter estimation represents the 3-D NLS approach, which is not practically feasible due to its high computational complexity. A practical alternative is the maximization of the 3-D periodogram [31], which is given by the magnitude square of the 3-D finite discrete-time Fourier transform (DTFT)

$$X(\lambda, \mu, \nu) = \sum_{l_s=0}^{L_s-1} \sum_{m_s=0}^{M_s-1} \sum_{n_s=0}^{N_s-1} \dots w_\lambda(l_s) w_\mu(m_s) w_\nu(n_s) \times x(l_s, m_s, n_s) e^{-j(\lambda l_s + \mu m_s + \nu n_s)} \quad (2)$$

for $\lambda \in [0, 2\pi)$, $\mu \in [0, 2\pi)$, and $\nu \in [0, 2\pi)$. Herein, $w_\lambda(l_s)$, $w_\mu(m_s)$, and $w_\nu(n_s)$ are normalized window functions in range dimension, relative velocity dimension, and angular dimension, respectively. A frequency discretization, as calculated via the discrete Fourier transformation (DFT), is obtained by setting $\lambda = (2\pi/L)l$, $l = 0, \dots, L - 1$, $\mu = (2\pi/M)m$,

$m = 0, \dots, M - 1$, and $\nu = (2\pi/N)n, n = 0, \dots, N - 1$, where L , M , and N are the number of frequency samples in range dimension, relative velocity dimension, and angular dimension, respectively. Frequency parameter estimation is simply done by peak searching in the periodogram.

The estimation of frequencies via the periodogram depends on the number of targets: for $K = 1$ the periodogram maximizer and the NLS estimator are equal, whereas for $K > 1$, the K largest peaks of the periodogram approximate the NLS estimator if the frequency separation of all possible target pairs is larger than the periodogram's resolution limit in at least one dimension [31, Sec. 4.3]. The resolution limits in the range dimension, relative velocity dimension, and angular dimension are $2\pi/L_s$, $2\pi/M_s$, and $2\pi/N_s$, respectively. If the separation is below these resolution limits in all three dimensions simultaneously, the periodogram fails to resolve the corresponding frequencies.

Strictly speaking, the NLS approximation of the periodogram approach holds only for constant window functions. However, in practice, one wants to distinguish targets with significantly different powers in their radar return so that window functions, as in [32], have to be used to reduce leakage. The downside of this approach is an increased estimation error, and a tradeoff has to be found in practice [31, Sec. 4.9].

Resolution and ambiguity

The periodogram's frequency resolution limits are straightforwardly translated to range, relative velocity, and angular resolution limits, which are $c/(2B)$, $1/(2\kappa M_s t_r)$, $\arcsin 1/(\kappa N_s d)$, respectively, where B denotes the bandwidth and t_r the pulse repetition time.

A practical approach to increase the resolution in the relative velocity and angular dimension, without increasing the number of pulses and receive channels, is undersampling, i.e., using pulse repetition times [33, Sec. 4.3] or antenna spacings [18, Sec. 2.4], which violate the sampling theorem. In the spatial domain, this means using an antenna spacing larger than half the wavelength, which has the additional benefit of being more robust against mutual coupling [34]. Due to aliasing, undersampling leads to ambiguities in the relative velocity and angular frequency parameters, respectively.

Table 1. The system parameters of a series-production automotive radar sensor.

Parameter	Value
Center frequency	76.15 GHz
Pulse bandwidth	200 MHz
Pulse repetition time	89 μ s
ULA antenna distance	12.8 mm
Number of samples per pulse	512
Number of pulses	256
Number of ULA antennas	4

To determine the relative velocity unambiguously, the pulse repetition time can be varied on a CPI basis and detections can be associated in range and angle over two subsequent CPIs [33, Sec. 8.3]. Another approach is to consider multiple relative velocity hypotheses in target tracking [35]. For resolving angular ambiguities, a single additional antenna can be employed, which has a half wavelength distance to one of the ULA antennas. The unambiguous angle obtained from the phase difference of such an antenna pair is then used to determine the correct angular hypothesis of the ULA-based measurement. This is a common approach in multiple baseline interferometry [36].

Practical aspects

For the calculation the 3-D periodogram, efficient algorithms and hardware accelerated fast Fourier transform (FFT) processors are available, e.g., [37]. A calculation sequence, which is particularly advantageous with respect to storage space, is described in [38]. Here, it is not necessary to store the complete data cube (1) before preprocessing which amounts to 2 MB for the described radar system in Table 1. Instead, the FFT in range dimension is calculated for each pulse and antenna, during data acquisition, and the intermediate result is stored in a compressed format. Subsequently, the FFT in relative velocity dimension and angular dimension is calculated for each range gate. For further processing, the result is only stored for power-detected processing cells and adjacent processing cells.

To cope with time-varying noise and interference statistics, constant false alarm rate (CFAR) methods [7, Ch. 16] are used for power detection. In particular, ordered-statistic CFAR methods [39] can be efficiently implemented [40].

For each range gate, peak frequencies in two dimensions are obtained from detected processing cells. An association over multiple range gates yields peak frequencies in all three dimensions. The 3-D peak frequencies represent the periodogram estimates discussed in the previous section, where the number of detected 3-D peak frequencies serves as the number of targets K .

Frequency estimates can be refined to subgrid accuracy by using frequency samples in the vicinity of the periodogram peak. For this, a wealth of methods exists, where an overview is given in [41]. However, the majority of methods restrict either the choice of window functions or the frequency grid size. For arbitrary window functions and grid sizes, a popular approach is the quadratic peak interpolation, which determines the maximizing frequency of an interpolated parabola, using the periodogram peak value and the two adjacent values. As an alternative, we use a simple, yet accurate look-up table approach, which determines the maximizing frequency from a look-up table, depending on the ratio of the periodogram peak value and the larger adjacent value. The look-up table can be calculated offline for the selected window function and the required subgrid accuracy. To this end, the true maximizing frequency is varied within one grid step, and the resulting periodogram ratio is inverted and linearly interpolated.

Automotive use case

In most automotive scenarios, the targets are well separated, i.e., their associated frequency differences exceed the periodogram's resolution limit in at least one dimension. However, specular multipath propagation can give rise to target pairs with frequency separation below the resolution limits in all three dimensions. In those cases, conventional processing fails, and high-resolution techniques become necessary. We give an example of such a multipath scenario, which constitutes a typical ACC or FCA use case.

Figure 1(a) shows a subject vehicle overtaking a slower target vehicle on a two-lane highway in a country where driving on the right is the norm. In this setup, the radar receives not only a direct target return, but also an indirect target return via the guard rail. The corresponding propagation paths are shown as dashed lines in Figure 1(a). This phenomenon is known as *specular multipath propagation* and leads to a mirror target in the radar return.

For calculating the range, relative velocity, and angular separation between the original target and the mirror target, we consider a lane width of 3.75 m, a center guard rail, a subject vehicle speed of 100 km/h, and a target vehicle speed of 80 km/h. We map the parameter separation to normalized frequency separation using the system parameters, gathered in Table 1, of a typical series-production automotive radar sensor. Figure 1(b) shows the frequency separations normalized to the respective resolution limit over the relative x -position of the target. The shape of the angular separation stems from undersampling in the angular dimension, as multiple angular hypotheses for the original and mirror target have to be considered.

Observe that above 130 m and between 25 m and 42 m, the frequency separation is below the respective resolution limit in all three dimensions, so that the periodogram will fail to resolve the original and mirror target. This can lead to misplaced target estimates in the driving path of the subject vehicle and may trigger erroneous ACC or FCA reactions such as deceleration or even emergency braking. This holds in particular for small x -positions in the region from 25 m to 42 m. To cope with such scenarios, we next present a framework for high-resolution frequency estimation.

High-resolution processing

High-resolution frequency estimation is theoretically able to resolve multiple targets if the frequencies are distinct in at least one dimension. Other works in automotive radar typically focus on high-resolution frequency estimation in the angular dimension. We point out that in critical use cases, as described in the section "Automotive Use Case," it can be advantageous to also consider the range and relative velocity dimension for high-resolution processing. In the following, we develop a framework to exploit the frequency separation in all three dimensions.

An optimal approach for high-resolution frequency estimation is to estimate all three frequency dimensions jointly, which is computationally demanding due to the large parameter space. In automotive radar, computational efficiency is crucial to ensure fast system reactions for ADASs and for future HAD

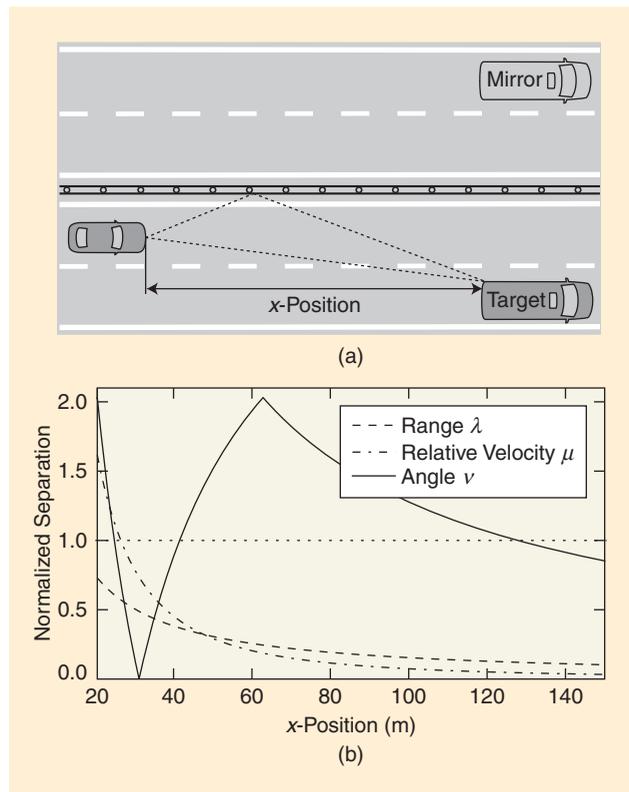


FIGURE 1. The two-target example: (a) the practically relevant scenario for ACC or FCA and (b) a corresponding normalized frequency separation. For the mapping of range, relative velocity, and angle to normalized frequencies, the radar system parameters in Table 1 are used.

applications. A common approach to reduce the computational complexity is to decouple the multidimensional frequency estimation into a sequence of 1-D frequency estimation problems. Note that a key result of [43] and [44] is that the decoupled approach can achieve almost the same estimation performance but with a significant reduction in computational cost. When decoupling the 3-D frequency estimation, one has to decide on the processing sequence. In one dimension, referred to as *resolution dimension*, a 1-D high-resolution frequency estimation of K targets has to be performed first. In the remaining dimensions, the calculated frequency estimates can then be used for signal component extraction, so that the remaining estimation problem is further simplified into K single-target frequency estimation problems. Note that the computational cost is dominated by the 1-D high-resolution frequency estimation of K targets in the resolution dimension. The subsequent single-target frequency estimators can be calculated by an computationally efficient periodogram-like approach.

The overall success of a decoupled approach depends critically on resolved estimates in the resolution dimension. It is well known that the resolution success of 1-D high-resolution frequency estimation depends on the available signal-to-noise ratio and particularly on the frequency separation. This holds for parametric approaches methods [18] as well as for compressed sensing approaches [20]. Therefore, the correct selection of the resolution dimension is crucial for

decoupled frequency estimation, where the best results are achieved when it is selected according to the largest frequency separation.

As mentioned in the section “Practical Aspects,” high-resolution processing in automotive radar is not only constrained by limited computational resources but also by memory. When applied in the original domain the complete data cube has to be stored, whereas Fourier-domain processing enables a memory-efficient implementation by storing only a reduced number of processing cells around power detections. Decoupled frequency estimation has been applied in [43] and [44] in the original domain. We extend this approach by applying decoupled frequency estimation in the Fourier domain. Note that this is particularly adapted to automotive radar use cases, in which the number of well-separated targets is typically very large, e.g., several hundreds. Fourier-domain processing allows a subdivision into smaller frequency estimation problems with significantly fewer targets. In most cases, only a single target is presented in a Fourier-domain processing cell.

In the sequel, we consider a framework for decoupled high-resolution frequency estimation in the Fourier domain, which allows a flexible selection of the resolution dimension and distinguishes a single target from multiple target processing cells. A special case for high-resolution frequency estimation in the spatial domain is also presented, corresponding to the considered DOA estimation problem in [25]. This approach has to be applied when the array elements are not uniformly spaced and can be advantageous when the number of array elements is very small so that element-space algorithms are computationally more efficient.

Decoupled models

Fourier domain

In automotive radar, a memory-efficient realization of decoupled high-resolution frequency estimation is based on a Fourier-domain model. After the 3-D finite DTFT calculation in (2), the model in (1) becomes

$$X(\lambda, \mu, \nu) = \sum_{k=1}^K a_k W_\lambda(\lambda - \lambda_k) \times W_\mu(\mu - \mu_k) W_\nu(\nu - \nu_k) + \Xi(\lambda, \mu, \nu), \quad (3)$$

where $W_\lambda(\lambda)$, $W_\mu(\mu)$, and $W_\nu(\nu)$ are the 1-D finite DTFT of window functions $w_\lambda(l_s)$, $w_\mu(m_s)$, and $w_\nu(n_s)$, respectively, and $\Xi(\lambda, \mu, \nu)$ is the 3-D finite DTFT of $\xi(l_s, m_s, n_s)$. Note that $\Xi(\lambda, \mu, \nu)$ is colored, circular complex Gaussian noise. However, for the purpose of frequency estimation, it can be assumed to be approximately white [42].

For decoupled high-resolution processing, the resolution dimension can be either the range dimension, the velocity dimension, or the angular dimension of the model in (3). A vectorization in the resolution dimensions enables three local vector models

$$z_\lambda(\mu, \nu) = \sum_{k=1}^{K_r} f_k(\mu, \nu) \mathbf{w}(\lambda_k) + \text{noise} \quad (4)$$

$$z_\mu(\lambda, \nu) = \sum_{k=1}^{K_r} f_k(\lambda, \nu) \mathbf{w}(\mu_k) + \text{noise} \quad (5)$$

$$z_\nu(\lambda, \mu) = \sum_{k=1}^{K_r} f_k(\lambda, \mu) \mathbf{w}(\nu_k) + \text{noise}. \quad (6)$$

Herein,

$$f_k(\mu, \nu) = a_k W_\mu(\mu - \mu_k) W_\nu(\nu - \nu_k) \quad (7)$$

$$f_k(\lambda, \nu) = a_k W_\lambda(\lambda - \lambda_k) W_\nu(\nu - \nu_k) \quad (8)$$

$$f_k(\lambda, \mu) = a_k W_\lambda(\lambda - \lambda_k) W_\mu(\mu - \mu_k) \quad (9)$$

gather the model terms in the respective remaining dimensions, and

$$\mathbf{w}(\lambda_k) = \left[W_\lambda\left(\frac{2\pi}{L}l_a - \lambda_k\right), \dots, W_\lambda\left(\frac{2\pi}{L}l_b - \lambda_k\right) \right]^T \quad (10)$$

$$\mathbf{w}(\mu_k) = \left[W_\mu\left(\frac{2\pi}{M}m_a - \mu_k\right), \dots, W_\mu\left(\frac{2\pi}{M}m_b - \mu_k\right) \right]^T \quad (11)$$

$$\mathbf{w}(\nu_k) = \left[W_\nu\left(\frac{2\pi}{N}n_a - \nu_k\right), \dots, W_\nu\left(\frac{2\pi}{N}n_b - \nu_k\right) \right]^T \quad (12)$$

are the model vectors in the respective resolution dimension, where $l_a, l_b, m_a, m_b, n_a,$ and n_b are chosen such that indices $l = l_a, \dots, l_b, m = m_a, \dots, m_b,$ and $n = n_a, \dots, n_b,$ respectively, contain the support for K_r local targets of interest. Typically, in an unresolved situation, this corresponds to one or two samples around the detected local maximum. Note that for the local models, we consider a reduced number of frequencies $K_r \ll K$ in the vicinity of detections. In particular, for most automotive radar scenarios, we use either $K_r = 1$ or $K_r = 2$ [25].

Spatial domain

A calculation similar to (2), but without the angular finite DTFT is

$$Y(\lambda, \mu; n_s) = \sum_{l_s=0}^{L_s-1} \sum_{m_s=0}^{M_s-1} w_\lambda(l_s) w_\mu(m_s) x(l_s, m_s, n_s) e^{-j(\lambda l_s + \mu m_s)} \quad (13)$$

with $\lambda \in [0, 2\pi)$, $\mu \in [0, 2\pi)$, and $n_s = 0, \dots, N_s - 1$. This form is required when the high-resolution processing is applied to the spatial domain and has been used in [25].

The signal model in (1) after 2-D finite DTFT calculation in (13) can be obtained similarly to (3). The corresponding local vector model is

$$\mathbf{y}(\lambda, \mu) = \sum_{k=1}^{K_r} f_k(\lambda, \mu) \mathbf{v}(\nu_k) + \text{noise}, \quad (14)$$

where $f_k(\lambda, \mu)$ is given in (9), and $\mathbf{v}(\nu_k) = [1, e^{j\nu_k}, \dots, e^{j(N_s-1)\nu_k}]^T$ is a ULA steering vector. Note that the model in (14) is in the original domain, in which the vector elements correspond to spatial array elements, whereas the model in (6) is in the Fourier domain, in which the vector elements correspond to samples of the angular spectrum.

Processing sequence

A maximum likelihood framework

For the Fourier domain models in (4)–(6), and the spatial domain model in (14), we want to decide between local signal models with $K_r = 1$ and $K_r = 2$, and estimate the respective unknown parameters. In a maximum likelihood framework, an optimal solution for this task is a generalized likelihood ratio test (GLRT) [45]. The GLRT statistic is given by the ratio of respective likelihood functions, which have been maximized with respect to the unknown parameters. A simplified version of the GLRT statistic can be obtained by the ratio of the mean squared errors of the respective models using the maximum likelihood parameter estimates.

The maximum likelihood framework is presented for the model in (4) only. It can be obtained accordingly for the models in (5), (6), and (14). Using the simplified version of the GLRT, a decision for a two-target situation is made if

$$\frac{\text{MSE}_{\lambda,1}}{\text{MSE}_{\lambda,2}} > \gamma,$$

where

$$\text{MSE}_{\lambda,1} = \frac{1}{l_b - l_a + 1} \|\mathbf{z}_\lambda - \mathbf{w}(\hat{\lambda}_0)\hat{f}_0\|^2 \quad (15)$$

$$\text{MSE}_{\lambda,2} = \frac{1}{l_b - l_a + 1} \|\mathbf{z}_\lambda - \mathbf{w}(\hat{\lambda}_1)\hat{f}_1 - \mathbf{w}(\hat{\lambda}_2)\hat{f}_2\|^2 \quad (16)$$

are the mean squared errors of (4) with $K_r = 1$ and $K_r = 2$, respectively, and γ is a suitable threshold. Herein, $\hat{\lambda}_0$ and \hat{f}_0 are the maximum likelihood estimates in the single target case, and $\hat{\lambda}_1$, \hat{f}_1 , $\hat{\lambda}_2$, and \hat{f}_2 are the maximum likelihood estimates for the two-target case.

The mean square errors corresponding to the models in (5), (6), and (14), are denoted $\text{MSE}_{\mu,1}$, $\text{MSE}_{\mu,2}$, $\text{MSE}_{\nu,1}$, $\text{MSE}_{\nu,2}$, MSE_1 , and MSE_2 , respectively, and can be obtained accordingly.

A suitable threshold γ can be obtained by fixing the false alarm rate to a certain level, where false alarm refers to the erroneous decision for the two-target case when only a single target is present. This can be done empirically via simulations and should be performed in a conservative way, such that the two-target case is only detected when reliable parameter estimation is possible.

The maximum likelihood estimates in the single target case are approximated using a look-up table or a quadratic interpolation approach around the maximum in the periodogram, as suggested in the section “Practical Aspects.” The required calculation is simple and can typically be performed for every detected processing cell. The maximum likelihood estimates in the two-target case are described in the section “High-Resolution Algorithms.” Here, the required calculation is computationally intensive and can only be performed for a selected subset of detected processing cells. This selection should take into account the deviation from the single target model and is described next.

Multiple target indication

A realization of the described maximum likelihood framework can be obtained by calculating the two-target maximum likelihood estimates only when the single target situation is unlikely and a multiple target situation is indicated. This indication is based on a goodness-of-fit test of the single target model [25], [46]. A test with low computational cost is given by

$$\text{MSE}_{\lambda,1} > T,$$

where T is a suitable threshold that depends on the noise power. In simple words, a single target situation is considered if $\text{MSE}_{\lambda,1}$ is of similar magnitude as the estimated noise power, which can be estimated from neighboring processing cells without targets. A multiple target indication is considered if $\text{MSE}_{\lambda,1}$ is significantly larger than the estimated noise power. A suitable threshold T can be obtained by fixing the false alarm rate to a certain level. This can be done empirically via simulations, or using the approximate distribution of the test statistic under the single target model. Further practical considerations, taking into account model deviations due to a weak secondary target or an imperfectly calibrated array, are described in [25]. Note, that the threshold T depends not only on the noise power but also on the sample support and applied window function in the resolution dimension. However, we omit this dependency for notational simplicity.

Optimal selection of the resolution dimension

The performance of high-resolution frequency estimation mainly depends on the available signal-to-noise ratio and the frequency separation normalized to the available sample support. For decoupled multidimensional frequency estimation, the best result will thus be obtained when the dimension with the largest frequency separation is selected as the resolution dimension. We have found in simulations that this corresponds to the largest mean squared error of the single target model, so that the resolution dimension is selected according to the largest value among $\text{MSE}_{\lambda,1}$, $\text{MSE}_{\mu,1}$, and $\text{MSE}_{\nu,1}$.

Overview

Figure 2(a) shows the signal flowchart for a flexible framework of high-resolution frequency estimation in the Fourier domain with optimal selection of resolution dimension. Figure 2(b) shows the flowchart for high-resolution frequency estimation in the spatial domain, corresponding to the proposed approach in [25], in which the resolution dimension is fixed to the spatial dimension.

High-resolution algorithms

Resolution dimension

The decoupled parameter estimation method is presented for the model in (4) only. It can be obtained accordingly for the models in (5), (6), and (14). In this section, we consider the model in (4) with $K_r = 2$

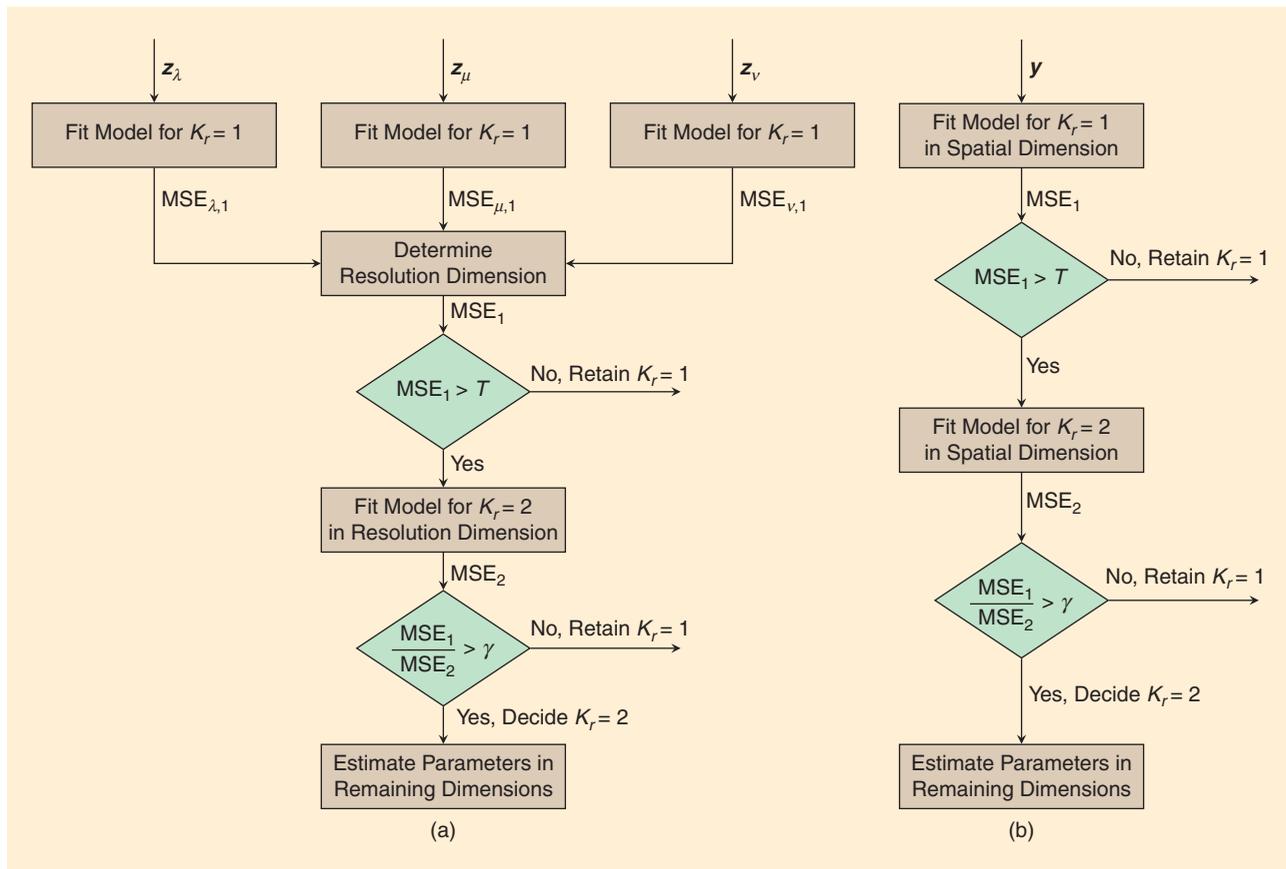


FIGURE 2. The flowchart overview: (a) high-resolution frequency estimation in the Fourier domain with an optimal selection of resolution dimension and (b) high-resolution frequency estimation in the spatial domain.

$$z_\lambda(\mu, \nu) = \mathbf{W}(\lambda_1, \lambda_2) [f_1(\mu, \nu), f_2(\mu, \nu)]^T + \text{noise}, \quad (17)$$

where $\mathbf{W}(\lambda_1, \lambda_2) = [\mathbf{w}(\lambda_1), \mathbf{w}(\lambda_2)]$.

The NLS optimization problem for $K_r = 2$ consists of finding frequencies λ_1 and λ_2 , which maximize criterion function

$$z_\lambda^H \mathbf{P}(\lambda_1, \lambda_2) z_\lambda, \quad (18)$$

where $\mathbf{P}(\lambda_1, \lambda_2) = \mathbf{W}(\lambda_1, \lambda_2) \mathbf{W}(\lambda_1, \lambda_2)^+$ is the projection matrix onto the column span of $\mathbf{W}(\lambda_1, \lambda_2)$, and $(\cdot)^+$ denotes the Moore–Penrose pseudoinverse operator. Note that we dropped the dependencies of μ and ν in (18) for convenience.

The NLS optimization is performed in two steps. In the first step, initial estimates for λ_1 and λ_2 are determined from a search on a coarse grid. The calculation of the criterion function (18) on a fixed grid can be simplified by considering techniques suggested in [25]. In particular, the rank-2 projection matrix $\mathbf{P}(\lambda_1, \lambda_2)$ can be eigendecomposed as $\mathbf{p}_1 \mathbf{p}_1^H + \mathbf{p}_2 \mathbf{p}_2^H$, and the projection operators \mathbf{p}_1 and \mathbf{p}_2 can be calculated offline for all required grid points. In this case, the calculation of the criterion function in (18) is effectively reduced to $|\mathbf{p}_1^H z_\lambda|^2 + |\mathbf{p}_2^H z_\lambda|^2$.

In the second step, the initial estimates for λ_1 and λ_2 are refined using few iterations of a Newton-type gradient search [45]. In each iteration, estimates are refined based on the gradient and the Hessian matrix of the criterion function

in (18) evaluated at the frequencies of the current iteration. The gradient can be determined directly using the derivative of the projection matrix, whereas the Hessian is approximated using the asymptotic expression given in [45]. As the Newton-type gradient search locally shows a quadratic convergence, often two iterations are sufficient to obtain the desired accuracy.

Regarding the computational complexity approximation of the described method, calculating the cost function on a coarse grid with L_g grid points requires $O(L_g^2 [l_b - l_a + 1]^2)$ operations and can be reduced to $O(L_g [l_b - l_a + 1])$ operations if the techniques in [25] are used. The Newton-type gradient search requires $O(K_r [l_b - l_a + 1]^2)$ operations per iteration [45]. In the implementation in the section “Experimental Results,” we use $L_g = 8$, $l_b - l_a + 1 = 5$ and limit the maximum number of iterations to four.

Alternative approaches for the iterative approximation of the maximum likelihood frequency estimator are the method of alternating projections [47] or the RELAX algorithm [48]. Note that both methods have also been applied in the context of automotive radar in [49] and [50], respectively.

Remaining dimensions

The frequencies in the remaining dimensions can be estimated by exploiting the structure of $f_1(\mu, \nu)$ and $f_2(\mu, \nu)$, given

Table 2. The decoupled parameter estimation overview.

Resolution Dimension			Remaining Dimensions		
Input	Algorithm	Output	Input	Algorithm	Output
\mathbf{z}_λ , (4)	NLS, $K_r = 2$	$\hat{\lambda}_1, \hat{\lambda}_2$	$\hat{f}_1(\mu, \nu), \hat{f}_2(\mu, \nu)$	NLS, $K_r = 1$	$\hat{\mu}_1, \hat{\mu}_2, \hat{\nu}_1, \hat{\nu}_2$
\mathbf{z}_μ , (5)	NLS, $K_r = 2$	$\hat{\mu}_1, \hat{\mu}_2$	$\hat{f}_1(\lambda, \nu), \hat{f}_2(\lambda, \nu)$	NLS, $K_r = 1$	$\hat{\lambda}_1, \hat{\lambda}_2, \hat{\nu}_1, \hat{\nu}_2$
\mathbf{z}_ν , (6)	NLS, $K_r = 2$	$\hat{\nu}_1, \hat{\nu}_2$	$\hat{f}_1(\lambda, \mu), \hat{f}_2(\lambda, \mu)$	NLS, $K_r = 1$	$\hat{\lambda}_1, \hat{\lambda}_2, \hat{\mu}_1, \hat{\mu}_2$
\mathbf{y} , (14)	NLS, $K_r = 2$	$\hat{\nu}_1, \hat{\nu}_2$	$\hat{f}_1(\lambda, \mu), \hat{f}_2(\lambda, \mu)$	NLS, $K_r = 1$	$\hat{\lambda}_1, \hat{\lambda}_2, \hat{\mu}_1, \hat{\mu}_2$

in (7). Considering the model in (17) with known λ_1 and λ_2 , estimates for $f_1(\mu, \nu)$ and $f_2(\mu, \nu)$, can be obtained using a linear least squares approach. By substituting estimates $\hat{\lambda}_1$ and $\hat{\lambda}_2$, we obtain

$$[\hat{f}_1(\mu, \nu), \hat{f}_2(\mu, \nu)]^T = \mathbf{W}(\hat{\lambda}_1, \hat{\lambda}_2)^+ \mathbf{z}_\lambda(\mu, \nu), \quad (19)$$

which is, in turn, used to estimate the frequencies in the remaining dimensions. In two NLS optimizations for $K_r = 1$, we estimate μ_1 and ν_1 by maximizing $|\hat{f}_1(\mu, \nu)|^2$, and we estimate μ_2 and ν_2 by maximizing $|\hat{f}_2(\mu, \nu)|^2$.

For calculating the estimates, (19) is used to evaluate $\hat{f}_1(\mu, \nu)$ and $\hat{f}_2(\mu, \nu)$ at a few DFT grid points around the initial parameter estimates, e.g., $m = m_a, \dots, m_b$ and $n = n_a, \dots, n_b$, the global maximum is determined, and a look-up table or a quadratic interpolation approach, as suggested in the section “Practical Aspects,” is used to refine the estimates. Given $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are sufficiently close to the respective true values, the calculated frequency estimates for the remaining dimensions approximate the maximum likelihood estimates [44]. Note that the computational cost of frequency estimation in the remaining dimensions is small when compared to the high-resolution frequency estimation in the resolution dimension.

The processing sequence of decoupled frequency estimation has been presented in detail for the case when the resolution dimension is λ , corresponding to the model in (4). It can be obtained accordingly for the models in (5), (6), and (14). An overview of decoupled parameter estimation is given in Table 2, which gathers input values, output values, and the employed algorithms. Note that the first row corresponds to the case when the resolution dimension is λ , whereas the resolution dimension is μ and ν in the second and third row, respectively. The last row corresponds to the special case of DOA estimation in the spatial domain, where the resolution dimension is fixed to the spatial dimension.

Experimental results

We use experimental data of a series-production automotive radar sensor to compare conventional radar processing, high-resolution frequency estimation in the spatial domain, and the proposed high-resolution frequency estimation in the Fourier domain with optimal selection of resolution dimension. Note that the radar system parameters are given in Table 1 and are

thus the same as the ones considered in the section “Automotive Use Case.”

The experimental setup is designed such that resolution is the dominant influence on the target position estimates and detections can be straightforwardly associated to the targets of interest. To this end, corner reflectors are employed to obtain point-target radar responses. The radar sensor is kept stationary, and the targets are moved toward the radar sensor to easily separate the targets from the environment, having a different relative velocity. Overall, this results in a clean two-target scenario without additional effects. Figure 3(a) shows the experimental setup. The two corner reflectors are placed on a carrier, which allows for variable relative positioning. For the two experiments the corner reflectors are positioned such that either the angular or the range dimension

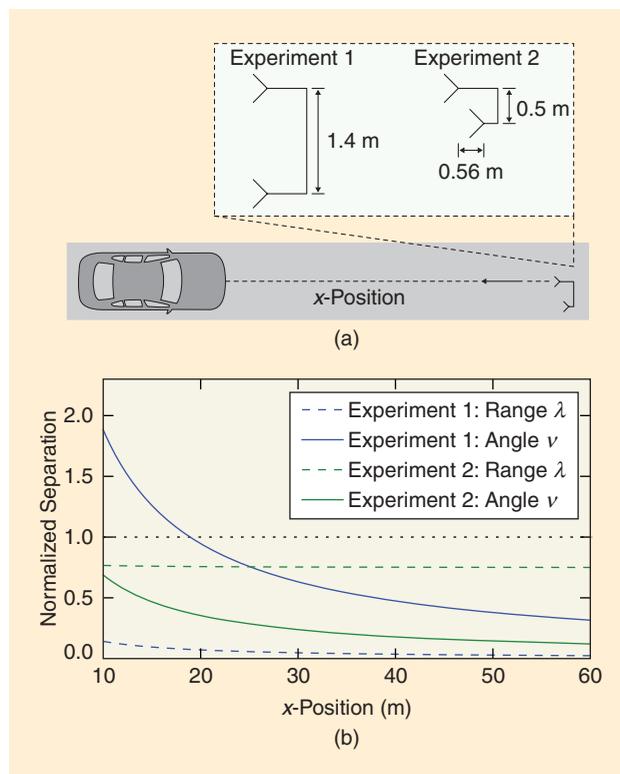


FIGURE 3. The experimental setup with (a) two corner reflectors and (b) the corresponding normalized frequency separation. For the mapping of range and angle to normalized frequencies, the radar system parameters in Table 1 are used.

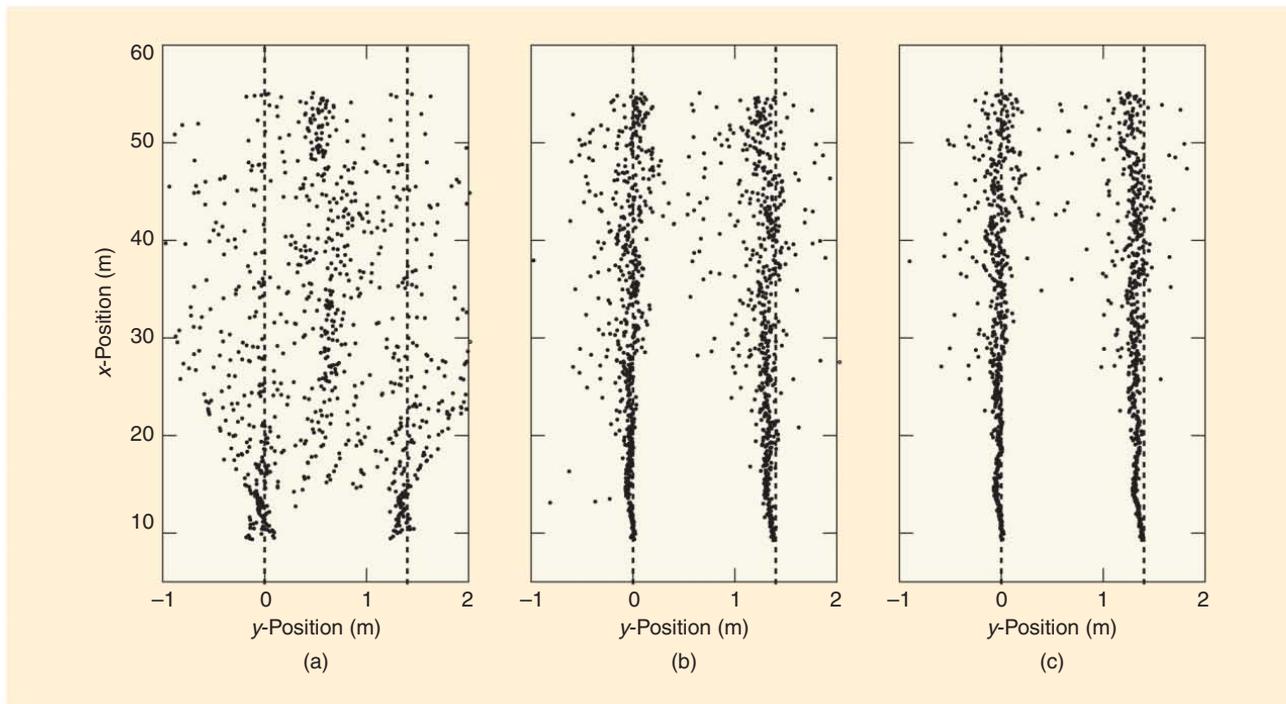


FIGURE 4. Real data results of a series-production automotive radar sensor for experiment 1 as described in Figure 3. Position estimates in Cartesian coordinates were obtained with (a) conventional radar processing, (b) high-resolution frequency estimation in the spatial domain, and (c) high-resolution frequency estimation in the Fourier domain with optimal selection of resolution dimension.

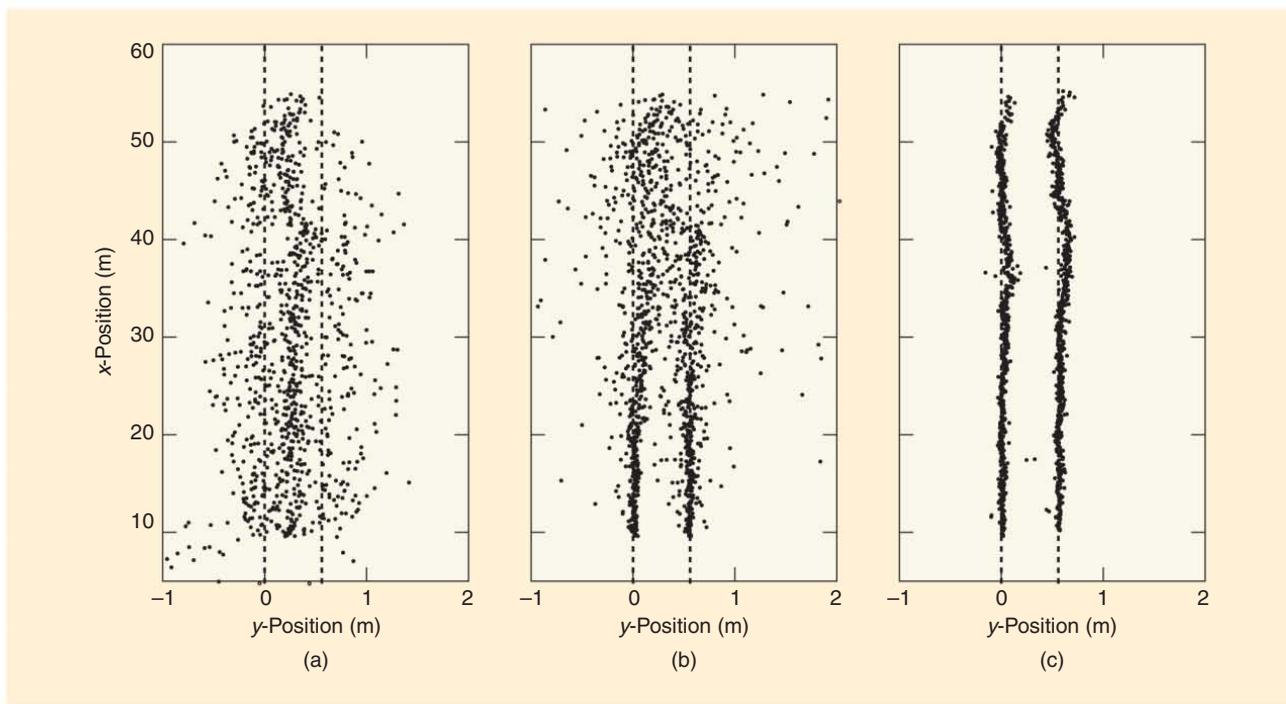


FIGURE 5. Real data results of a series-production automotive radar sensor for experiment 2 as described in Figure 3. Position estimates in Cartesian coordinates were obtained with (a) conventional radar processing, (b) high-resolution frequency estimation in the spatial domain, and (c) high-resolution frequency estimation in the Fourier domain with optimal selection of resolution dimension.

has a significantly larger frequency separation than all others. The carrier is then moved straight toward the radar sensor at walking speed.

Figure 3(b) shows the corresponding frequency separations normalized to the respective resolution limits. The relative velocity separation is very low and thus omitted. Observe that,

for the first experiment, the frequency separation in the angular dimension is the largest, whereas for the second experiment the frequency separation in the range dimension is the largest. For all dimensions, the separations are below the resolution limit except for x -positions below 18 m, where the angular separation exceeds the resolution limit for the first experiment. Therefore, we expect that conventional radar processing will fail for both experiments, except for x -positions below 18 m in the first experiment. For the first experiment, we further expect that high-resolution processing in the spatial domain will perform well due to the dominating angular separation and that high-resolution processing in the Fourier domain selects the angular dimension as resolution dimension, leading to a similar performance. For the second experiment, we expect that high-resolution processing in the Fourier domain correctly selects the range dimension as resolution dimension and will thus provide resolved and accurate frequency estimates. In contrast, we expect that high-resolution processing in the spatial domain performs significantly worse due to the predefined resolution dimension and the much smaller angular separation.

Figure 4 shows the resulting position estimates in Cartesian coordinates for the first experiment, which are obtained with (a) conventional radar processing, (b) high-resolution frequency estimation in the spatial domain, and (c) high-resolution frequency estimation in the Fourier domain with optimal selection of resolution dimension. The true target positions are shown as dashed vertical lines. As expected, conventional processing fails to resolve the two targets except for x -positions below 18 m. In contrast, both high-resolution approaches are able to resolve the targets and provide accurate position estimates. Note that this means that the resolution dimension was correctly selected by the approach proposed in the section “Optimal Selection of the Resolution Dimension.”

Figure 5 shows the corresponding results for the second experiment. Here, conventional processing fails for all x -positions. The high-resolution approach in the spatial domain breaks down for x -positions above 30 m due to the small frequency separation in the angular dimension. Only the high-resolution, Fourier-domain approach can resolve the two targets for all x -positions, as it exploits the much larger range separation. Note that this is enabled by the approach proposed in the section “Optimal Selection of the Resolution Dimension,” which correctly selects the range dimension as the resolution dimension.

Conclusions

We have considered automotive radar as a key technology for ADASs and HAD. Current limitations with respect to the requirements of the automotive industry can be met by using high-resolution frequency estimation. We have described conventional automotive radar processing and pointed out use cases in which it is bound to fail. We have presented a flexible framework for computationally efficient high-resolution frequency estimation as an enhancement to conventional radar processing. Real data from a series-production automotive radar sensor have been presented to show the effectiveness of the presented approach.

Authors

Florian Engels (florian.engels@continental-corporation.com) received his M.Sc. and Dr.-Ing. degrees, both in electrical engineering, from Karlsruhe Institute of Technology [formerly Universität Karlsruhe (TH)], Germany, in 2008 and 2016, respectively. Since 2008, he has been with Continental, Adaptive Distance Control Systems GmbH, Lindau, Germany, as a radar algorithm development engineer and, since 2015, as a leading expert for radar signal processing. His research interests include multidimensional frequency estimation with applications in automotive radar.

Philipp Heidenreich (philipp.heidenreich@opel.com) received his Dipl.-Ing. and Dr.-Ing. degrees, both in electrical engineering, from Technische Universität Darmstadt, Germany, in 2006 and 2012, respectively. From 2007 to 2012, he was with the Signal Processing Group, Technische Universität Darmstadt, as a research scientist. From 2012 to 2015, he was with Continental, Adaptive Distance Control Systems GmbH, Lindau, Germany, as a radar algorithm development engineer. Since 2015, he has been with General Motors Europe, Adam Opel AG, Rüsselsheim, Germany, as a systems engineer for automotive active safety applications. His research interests include statistical and array signal processing with applications to automotive radar. He received the Best Student Paper Award at the Fifth Biennial Workshop on Digital Signal Processing for In-Vehicle Systems in 2011, in Kiel, Germany. He is a Member of the IEEE.

Abdelhak M. Zoubir (zoubir@spg.tu-darmstadt.de) received his Dr.-Ing. degree from Ruhr-Universität Bochum, Germany, in 1992. He was with Queensland University of Technology, Australia, from 1992 to 1998, where he was an associate professor. In 1999, he joined Curtin University of Technology, Australia, as a professor of telecommunications. In 2003, he moved to Technische Universität Darmstadt, Germany, as a professor of signal processing and the head of the Signal Processing Group. His research interest lies in statistical methods for signal processing with emphasis on bootstrap techniques, robust detection and estimation, and array processing applied to telecommunications, radar, sonar, automotive monitoring and safety, and biomedicine. He has served on publication boards of various publications, most notably as the editor-in-chief of *IEEE Signal Processing Magazine* (2012–2014). He is a Fellow of the IEEE and is an IEEE Distinguished Lecturer (class of 2010–2011).

Friedrich K. Jondral (Friedrich.Jondral@kit.edu) received the diploma in mathematics and the doctoral degree in natural sciences from the Technische Universität Braunschweig, Germany, in 1975 and 1979, respectively. Since 1993, he has been a full professor and the director of the Communications Engineering Laboratory at the Karlsruhe Institute of Technology [formerly Universität Karlsruhe (TH)], Germany, where he served as dean of the Department of Electrical Engineering and Information Technology from 2000 to 2002. His current research interests include software-defined and cognitive radio, cooperation and coordination, signal analysis, pattern recognition, network capacity optimization, and dynamic

channel allocation. He has been a member of the Verband der Elektrotechnik Elektronik Informationstechnik e.V. since 1988. He is a Senior Member of the IEEE.

Markus Wintermantel (markus.wintermantel@continental-corporation.com) received his Dipl.-Ing. and Dr.-Ing. degrees, both in electrical engineering, from Universität Stuttgart, Germany, in 1992 and 2000, respectively. From 1992 to 2000, he was with the Institute of Network and System Theory, Universität Stuttgart, as a research scientist. Since 2000, he has been with Continental, Adaptive Distance Control Systems GmbH, Lindau, Germany. Currently, he is a senior expert for automotive radar sensors. He holds more than 20 patents in the area of automotive radar applications.

References

- [1] F. Gustafsson, "Automotive safety systems," *IEEE Signal Process. Mag.*, vol. 26, no. 4, pp. 32–47, July 2009.
- [2] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner, "Three decades of driver assistance systems: Review and future perspectives," *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 4, pp. 6–22, Oct. 2014.
- [3] A. Eskandarian, *Handbook of Intelligent Vehicles*. New York: Springer-Verlag, 2012.
- [4] M. Murad, I. Bilik, M. Friesen, J. Nickolaou, J. Salinger, K. Geary, and J. Colburn, "Requirements for next generation automotive radars," in *Proc. IEEE Radar Conf.*, Ottawa, Canada, May 2013.
- [5] J. Dickmann, N. Appenrodt, H. L. Bloecher, C. Brenk, T. Hackbarth, M. Hahn, J. Klappstein, M. Muntzinger, and A. Sailer, "Radar contribution to highly automated driving," in *Proc. European Radar Conf.*, Rome, Italy, Oct. 2014, pp. 412–415.
- [6] C. Waldschmidt and H. Meinel, "Future trends and directions in radar concerning the application for autonomous driving," in *Proc. European Radar Conf.*, Rome, Italy, Oct. 2014, pp. 416–419.
- [7] M. Richards, J. Scheer, and W. Holm, *Principles of Modern Radar Basic Principles*. Raleigh, NC: SciTech, 2010.
- [8] F. Fölster, "Erfassung ausgedehnter Objekte durch ein Automobil-Radar," Ph.D. dissertation, Universität Hamburg-Harburg, Dec. 2006.
- [9] M. Bühren, "Simulation und Verarbeitung von Radarziellisten im Automobil," Ph.D. dissertation, Univ. Stuttgart, Jan. 2008.
- [10] L. Hammarstrand, L. Svensson, F. Sandblom, and J. Sorstedt, "Extended object tracking using a radar resolution model," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 3, pp. 2371–2386, July 2012.
- [11] F. Diewald, J. Klappstein, F. Sarholz, J. Dickmann, and K. Dietmayer, "Radar-interference-based bridge identification for collision avoidance systems," in *Proc. IEEE Intelligent Vehicles Symp.*, Baden-Baden, Germany, June 2011, pp. 113–118.
- [12] S. Heuel, "Fußgängererkennung im Straßenverkehr mit 24GHz Radarsensoren," Ph.D. dissertation, Technische Univ. Hamburg-Harburg, Dec. 2013.
- [13] C. Lundquist, L. Hammarstrand, and F. Gustafsson, "Road intensity based mapping using radar measurements with a probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1397–1408, Apr. 2011.
- [14] C. Lundquist, T. Schön, and F. Gustafsson, "Situational awareness and road prediction for trajectory control applications," in *Handbook of Intelligent Vehicles*. New York: Springer-Verlag, 2012, pp. 366–396.
- [15] M. Clark and L. Scharf, "Two-dimensional modal analysis based on maximum likelihood," *IEEE Trans. Signal Process.*, vol. 42, no. 6, pp. 1443–1452, June 1994.
- [16] M. Haardt and J. Nosske, "Simultaneous Schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 46, no. 1, pp. 161–169, Jan. 1998.
- [17] M. Pesavento, "Fast algorithms for multidimensional harmonic retrieval," Ph.D. dissertation, Ruhr-Univ. Bochum, 2005.
- [18] H. van Trees, *Detection, Estimation, and Modulation Theory—Part IV Optimum Array Processing*. Hoboken, NJ: Wiley, 2002.
- [19] H. Krim and M. Viberg, "Two decades of array signal processing research," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, July 1996.
- [20] J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.*, vol. 67, no. 6, pp. 906–956, June 2014.
- [21] S. Fortunati, R. Grasso, F. Gini, M. Greco, and K. LePage, "Single-snapshot DOA estimation by using compressed sensing," *EURASIP J. Adv. Signal Process.*, June 2014.
- [22] P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 573–586, Feb. 2011.
- [23] P. Wenig, M. Schoor, O. Gunter, B. Yang, and R. Weigel, "System design of a 77 GHz automotive radar sensor with superresolution DOA estimation," in *Proc. Int. Symp. Signals, Systems and Electronics*, Montreal, Canada, July 2007, pp. 537–540.
- [24] M. Schoor, "Hochauflösende Winkelschätzung für automobile Radarsysteme," Ph.D. dissertation, Univ. Stuttgart, 2010.
- [25] P. Heidenreich, "Antenna array processing: Autocalibration and fast high-resolution methods for automotive radar," Ph.D. dissertation, Technische Univ. Darmstadt, 2012.
- [26] H. Winner, "Radarsensorik," in *Handbuch Fahrerassistenzsysteme – Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort*. New York: Springer-Verlag, 2015, pp. 259–316.
- [27] M. Kronauge, "Waveform design for continuous wave radars," Ph.D. dissertation, Technische Universität Hamburg-Harburg, July 2014.
- [28] W. Menzel and A. Moebius, "Antenna concepts for millimeter-wave automotive radar sensors," *Proc. IEEE*, vol. 100, no. 7, pp. 2372–2379, July 2012.
- [29] F. Gini, A. Maio, and L. Patton, *Waveform Design and Diversity for Advanced Radar Systems*. London: IET, 2012.
- [30] M. Wintermantel, "Radar system with improved angle formation," German Patent WO 2010/000252 A2, 2010.
- [31] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Englewood Cliffs, NJ: Prentice Hall, 2005.
- [32] F. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proc. IEEE*, vol. 66, no. 1, pp. 51–83, Jan. 1978.
- [33] N. Levanon and E. Mozeson, *Radar Signals*. Hoboken, NJ: Wiley, 2004.
- [34] T. Svantesson, "Antennas and propagation from a signal processing perspective," Ph.D. dissertation, Chalmers Univ. Technol., Gothenburg, Sweden, 2001.
- [35] K. Li, B. Habtemariam, R. Tharmarasa, M. Pelletier, and T. Kirubarjan, "Multitarget tracking with Doppler ambiguity," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 4, pp. 2640–2656, Oct. 2013.
- [36] E. Jacobs and E. W. Ralston, "Ambiguity resolution in interferometry," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 17, no. 6, pp. 766–780, Nov. 1981.
- [37] D. Nugraha, A. Roger, and R. Ygnace, "Integrated FFT accelerator and inline bin-rejection for automotive FMCW radar signal processing," in *Proc. European Radar Conf.*, Paris, France, Sept. 2015, pp. 546–567.
- [38] M. Wintermantel, "Radar system with optimized storage of temporary data," German Patent WO 2015/185058 A1, 2015.
- [39] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, no. 4, pp. 608–621, July 1983.
- [40] M. Wintermantel, "New approach for an approximate realization of rank order filtering with low complexity," *Frequenz*, vol. 60, no. 1/2, pp. 24–28, Feb. 2006.
- [41] E. Jacobsen and P. Kootsookos, "Fast, accurate frequency estimators," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 123–125, May 2007.
- [42] P. Stoica, A. Jakobsson, and J. Li, "Cisoid parameter estimation in the colored noise case: Asymptotic Cramér-Rao bound, maximum likelihood, and nonlinear least-squares," *IEEE Trans. Signal Process.*, vol. 45, no. 8, pp. 2048–2059, Aug. 1997.
- [43] A. Swindlehurst and P. Stoica, "Maximum likelihood methods in radar array signal processing," *Proc. IEEE*, vol. 86, no. 2, pp. 421–441, Feb. 1998.
- [44] F. Athley, "Asymptotically decoupled angle-frequency estimation with sensor arrays," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, 1999, pp. 1098–1102.
- [45] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing," in *Radar Array Processing*. New York: Springer-Verlag, 1993.
- [46] P. Heidenreich, D. Stenmanns, and A. Zoubir, "Computationally simple criteria for detecting a multi-target scenario in automotive radar array processing," in *Proc. European Signal Processing Conf.*, Aalborg, Denmark, Aug. 2010, pp. 224–228.
- [47] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projections," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 36, no. 10, pp. 1553–1560, Oct. 1988.
- [48] J. Li and P. Stoica, "Efficient mixed-spectrum estimation with applications to target feature extraction," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 281–295, Feb. 1996.
- [49] P. Heidenreich and A. Zoubir, "High-resolution direction finding of coherent sources in the presence of model errors using alternating projections," in *Proc. IEEE Workshop Statistical Signal Processing*, Cardiff, U.K., Sept. 2009, pp. 521–524.
- [50] F. Engels, "Target shape estimation using an automotive radar," in *Smart Mobile In-Vehicle Systems—Next Generation Advancements*. New York: Springer-Verlag, 2014.

SP

Stefan Schwarz, Tal Filosof, and Markus Rupp

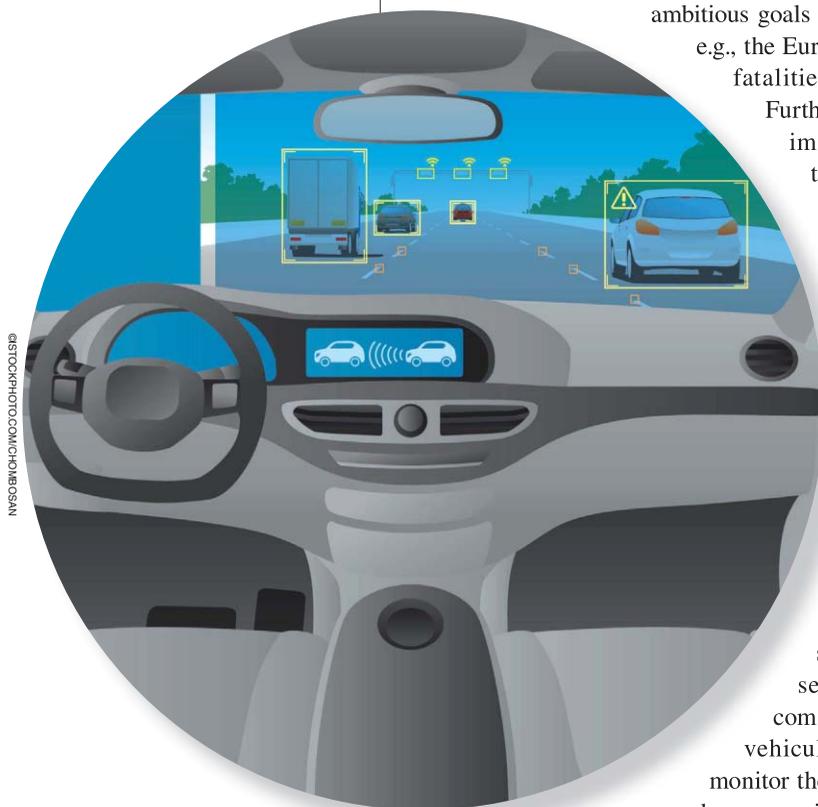
Signal Processing Challenges in Cellular-Assisted Vehicular Communications

Efforts and developments within 3GPP LTE and beyond

Vehicular communications is an important enabler for enhancing the safety on roads by supporting mutual awareness of vehicles as well as for improving the efficiency of transportation through smart traffic management by intelligent transport systems (ITSs). Governments around the world have set ambitious goals for road fatality reduction in the near future; e.g., the European Union targets a 50% reduction of road fatalities by 2020 as compared to the year 2010.

Furthermore, traffic telematic systems aim to minimize the environmental impact of transportation and maximize the utilization of available road infrastructure by adaptive traffic management. To realize these challenging targets, autonomous wireless information exchange among vehicles—vehicle to vehicle (V2V)—and with roadside infrastructure—vehicle to infrastructure (V2I)—are central ingredients. In addition to traffic efficiency and safety-related issues, vehicular communications is increasingly recognized as an important revenue driver by car manufacturing companies since it enables wirelessly connected in-vehicle entertainment systems that support on-demand video streaming and online Internet access for passengers. Also, in the future, machine-type communication is expected to play a major role in vehicular environments, with more sensors that monitor the internal state of vehicles and autonomously exchange service and maintenance information with cloud servers of manufacturers. Depending on the considered use-case,

distinct quality of service (QoS) requirements come into play [1]: infotainment applications for in-car users require high bandwidth and network capacity, active road safety relies on delay- and outage-critical data transmission, whereas information exchange for road traffic efficiency management typically comes without strict QoS requirements and exhibits graceful degradation of performance with increasing latency.



Digital Object Identifier 10.1109/MSP.2016.2637938
Date of publication: 3 March 2017

Introduction

In recent years, specific vehicular communication systems, such as the European Telecommunications Standards Institute (ETSI) ITS G5 based on IEEE 802.11p dedicated short-range communication (DSRC), have been developed to enable timely and reliable exchange of information in so-called vehicular ad hoc networks (VANETs). These systems are commonly based on dedicated infrastructure and employ licensed transmission bands to avoid interference with other existing systems. However, interest in mobile communications technology to support VANETs has been recently increasing, because this technology is available off-the-shelf and, therefore, enables cost-effective implementations. Due to the virtual ubiquity of cellular networks, vehicular communications strategies of many companies are shaped by the development of mobile network technologies. In response, the progression of the universal mobile telecommunications system (UMTS) long-term evolution (LTE) toward vehicular communications is currently pushed within the Third Generation Partnership Project (3GPP) to meet the requirements of vehicular environments.

Efficient and reliable wireless communication with users at high mobility, however, comes with several unique challenges that the LTE standard cannot yet stand up to, as we demonstrate in this article by some selected examples. Yet, advanced signal processing at the transmitters and the receivers has the potential to alleviate the shortcomings of LTE. In this article, we provide an overview of efforts ongoing within the development of LTE Release 14 to enhance the support of vehicular communications. We further give insights into promising signal processing methods for efficient wireless connectivity at high mobility. While some of the described techniques can readily be employed without modification of the LTE standard, others require additional standardization efforts to harmonize the operation of communication devices.

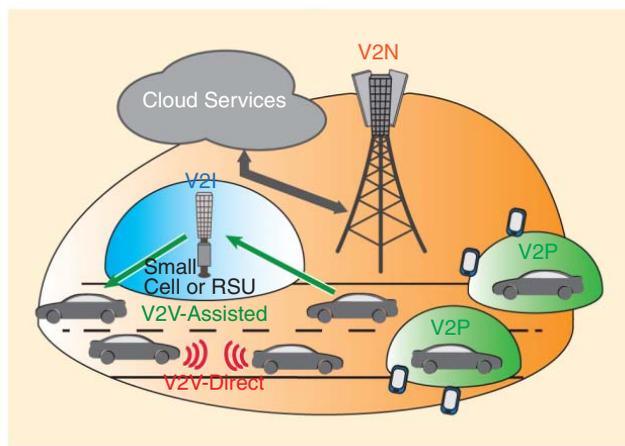


FIGURE 1. Vehicular communications scenarios considered within 3GPP: V2N provides connectivity to the cloud; V2I covers connections to RSUs; V2V deals with direct and assisted intervehicle transmission; V2P enables information exchange with other devices in the proximity of vehicles (e.g., pedestrians and cyclists).

LTE contenders for vehicular communications

The 3GPP recently initiated study items within Release 14 of UMTS LTE on vehicular communications; see [31] and [32]. The goal of these study items is to develop a set of LTE specifications for vehicular environments [LTE-based vehicle to X (V2X)] [2]. Currently, the focus of standardization is on exploring the support of active road safety applications, either in a dedicated transmission band or shared with conventional mobile subscribers. The different use cases considered within the 3GPP are illustrated in Figure 1. Connectivity in between vehicles, e.g., to enable the V2V exchange of location and trajectory information for driver assistance systems, is supported directly via device-to-device (D2D) transmission, as well as indirectly, employing base stations and dedicated roadside units (RSUs) as transmission hubs to facilitate information sharing over larger distances. The local exchange of traffic management information between ITS infrastructure and vehicles is covered under the umbrella of V2I communication. Connectivity to the Internet and to cloud services falls in the category of vehicle-to-network (V2N) transmission; finally, the exchange of mutual awareness information between vehicles and pedestrians (or cyclists) is supported by vehicle-to-pedestrian (V2P) transmission. Commonly, all of these different types of vehicular communication scenarios are summarized under the term V2X communication. Even though they appear very similar from a basic physical layer (PHY) perspective, the different use cases require individual treatment within standardization to enable efficient incorporation into the existing specifications [3].

Currently three technologies are considered as central for LTE based vehicular communications:

- *Dual connectivity to support high user mobility in dense heterogeneous networks (HetNets):* Mobile communication connections at high mobility suffer from frequent interruptions due to handovers in between cells; this is especially problematic in dense HetNets with coverage areas of base stations of mere tens to few hundreds of meters [4]. To alleviate such issues, the 3GPP has incorporated dual connectivity in LTE Release 12. With dual connectivity, robustness with respect to mobility is enhanced by connecting subscribers to two cells—master and secondary—simultaneously. Thereby, critical control and signaling information, as provided by the control plane of the mobile network, is kept in the master cell of the macro base-station layer, providing reliable coverage over large geographic areas with a minimal amount of handovers. The capacity and rate of the actual data transmission is enhanced by concurrently providing data over the user plane of the network from macro base stations and small cells [5]. For this data exchange, connection interruptions are less critical since they do not cause connectivity failures.
- *LTE-based broadcast services, as supported by the public warning system (PWS) and enhanced multimedia broadcast/multicast service (eMBMS), for efficient distribution of messages among vehicles:* In cooperative ITS (C-ITS), two basic types of messages are specified by ETSI for exchanging

information among vehicles: 1) The periodic cooperative awareness message (CAM) is employed to share vehicle status information (location, direction, speed) with cars in the vicinity, and 2) the event-driven sporadic decentralized environmental notification message is used to exchange critical warnings, e.g., in case of traffic accidents (in the U.S. Society of Automotive Engineers J2735 standard, both types are covered by the basic safety message). Since such messages are shared with all other vehicles within a certain geographic region, broadcasting/multicasting them over the cellular network can be much more efficient than unicast transmission to each vehicle individually. The distribution of messages in certain geographic notification areas is supported by the PWS, which allows consideration of geographic packet routing information. With Multimedia Broadcast/Multicast Single Frequency Network (MBSFN), the LTE standard even supports multicasting in potentially very large geographic areas by synchronously transmitting the same information from multiple base stations.

- *Proximity services, including D2D communication to realize connectivity in between vehicles as well as between connected cars and handheld terminals:* Especially in safety-critical situations, latency of data exchange can be a limiting factor. Lowest latency is achieved by minimizing the amount of traffic nodes that need to be traversed between source and destination; hence, direct D2D communication is most promising for short-range information exchange with stringent latency requirements. Furthermore, D2D transmission can be designed to enable autonomous VANET operation in areas that are not covered by cellular infrastructure, thus enhancing the reliability and availability of vehicular communications. Besides these operational advantages, D2D offloading of traffic from the cellular network also helps to reduce the overhead caused by vehicular communications.

Our focus in this article is on infrastructure-based vehicular communications, i.e., on V2I and cellular-assisted V2V transmission. We present exemplary performance results of LTE-compliant V2X transmissions to demonstrate important potential shortcomings of the standard that may limit the applicability of LTE for vehicular communications. Based on these insights, we then present signal processing techniques to mitigate these weaknesses.

Dual connectivity-enhanced V2I transmission

Dual connectivity is intended to solve the handover problem of subscribers moving at high mobility through dense HetNets. To investigate the performance of this approach, we conducted LTE-compliant system-level simulations of such dense HetNets, employing the Vienna LTE system-level simulator [6]; Figure 2 shows the macroscopic signal-to-interference-plus-noise ratio (SINR) distribution of one snapshot of the evaluated HetNet. The network contains seven hexagonally arranged macro base stations (BS 1,...,BS 7) that each serve three sectors, with nominal coverage regions illustrated as white hexagons. Additionally, a large number of small cells is randomly distributed over the network area; these devices are equipped with omnidirectional

antennas. Subscribers are served via dual connectivity from the strongest macro base station and small cell simultaneously, with the control plane being kept at the macro layer to minimize necessary handovers.

In our numerical investigation, we assume that all base stations operate with 5-MHz system bandwidth at the same carrier frequency $f_c = 2$ GHz, i.e., we consider a spatial frequency reuse-factor of one. Without dual connectivity, this leads to low SINR values at the cell edge as illustrated in Figure 2. Most contemporary fourth-generation cellular networks nevertheless utilize all available bandwidth at all base stations to enable harvesting the spatial reuse gain promised by network densification; this will hold even more for future fifth-generation (5G) networks. The considered situation corresponds to the so-called cochannel dual connectivity mode of LTE, where small cells and macro base stations employ the same carrier frequency for data transmission. Additionally, the LTE standard supports interfrequency dual connectivity, where different frequencies are employed on macro base stations and small cells. Notice that the SINR distribution in Figure 2 shows the behavior without cochannel dual connectivity. If dual connectivity is activated for a user, the cell edge between the corresponding macro base station and small cell vanishes, since data transmission to the user occurs from both base stations simultaneously. We evaluate the empirical cumulative distribution function (ECDF) of the downlink throughput of subscribers in Figure 3. We compare the performance of low- and high-mobility users moving at $v = 5$ km/h and $v = 150$ km/h, respectively. The system applies transmission rate adaptation according to the instantaneous SINR of subscribers, such as to match the spectral efficiency of the employed modulation and coding scheme (MCS) to the current channel quality. The required channel state information at the transmitter (CSIT) is provided as feedback information from the subscribers to the base stations using the LTE standard defined channel

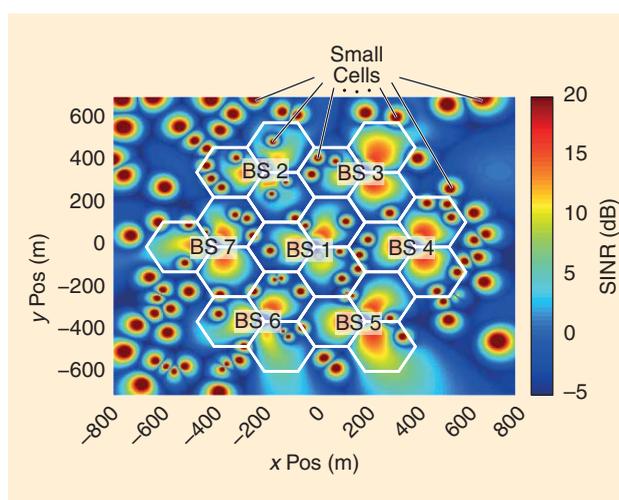


FIGURE 2. The macroscopic path loss-based SINR distribution of a random HetNet realization with seven macro base stations (each serving three sectors) and a small-cell density of 50 per km^2 (Poisson point process). The white hexagons illustrate the nominal assignment regions of macro base stations.

quality indicator (CQI) and the feedback algorithms described in [7]. More specifically, the CQI is utilized by the receiver to signal via a dedicated feedback link to the transmitter which MCS should be employed for transmission to achieve reliable as well as efficient transmission; this information is derived from the SINR currently experienced by the receiver. We observe from Figure 3 that, under ideal circumstances of delayless noncausal feedback (delay 0 ms), high-mobility users achieve almost the same performance as those at low mobility. There is only a small performance degradation, which is caused by the need to apply a slightly more conservative transmission rate adaptation, since the channel of high-mobility users varies even within one transmission time interval (TTI). This observation confirms the expected mobility enhancement achieved by dual connectivity. Yet, as soon as we consider certain delay in the CQI feedback link, we observe strong throughput degradation at high mobility, with more than 25% of subscribers obtaining zero throughput even with only 1 ms feedback delay. This performance degradation occurs because transmission rate adaptation with feedback delay is based on outdated channel state information (CSI) leading to signal outages due to mismatch between the utilized transmission rate and the rate supported by the channel. This loss in throughput goes hand in hand with an increase in latency, since each lost packet must be retransmitted. Notice, the TTI length of LTE is $T_s = 1$ ms. Hence, a delay of the feedback processing below this value is infeasible; even the value of 5 ms considered in Figure 3 may be hard to achieve in practice.

To gauge the expected impact of feedback delay τ in milliseconds, with respect to the system parameters carrier frequency f_c , TTI length T_s , and user velocity v , we define the normalized feedback delay as

$$\tau_n = \tau f_d = \tau_r T_s f_d = \tau_r T_s f_c \frac{v}{c}, \quad (1)$$

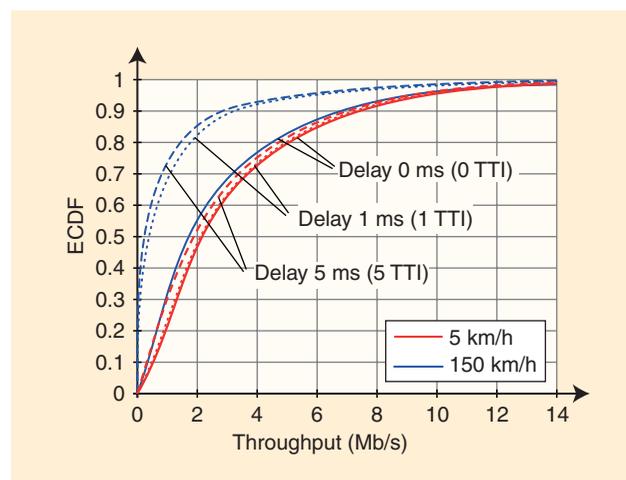


FIGURE 3. The throughput in dependence of user velocity and CSI feedback delay in a dense HetNet supporting dual connectivity.

CSI feedback delay is a significant issue for high-mobility users, and it can be a strongly limiting factor for the rate performance and reliability of wireless transmission in vehicular scenarios.

with f_d denoting the maximum Doppler frequency shift in hertz, $\tau_r = \tau/T_s$ being the relative feedback delay in multiples of TTIs, and c representing the speed of light. The feedback delay has to be seen in relation to the temporal variability of the channel; if the channel is quasi static (as in many indoor scenarios), a feedback delay is irrelevant. It only matters if the channel changes significantly within the duration of the feedback delay, which is gauged by

the normalized feedback delay. More specifically, the Doppler shift f_d is inversely proportional to the coherence time of the channel; hence, the product τf_d is proportional to the number of coherence intervals elapsing during the feedback delay duration. The exact proportionality constant depends on the observed Doppler spectrum [8]. In our aforementioned example, we have for $\tau = 1$ ms with $v = 5$ km/h and $v = 150$ km/h, $\tau_n \approx 0.009$, and $\tau_n \approx 0.27$, respectively. We will encounter the normalized feedback delay again later in the section “Future Enhancements and Challenges.”

CSI feedback delay is a significant issue for high-mobility users, and it can be a strongly limiting factor for the rate performance and reliability of wireless transmission in vehicular scenarios. The problem is, of course, not specific to dual connectivity, but occurs whenever rate adaptation is based on outdated CSI. In the section “CSI Feedback Enhancement,” we discuss possible signal processing approaches to mitigate this performance degradation.

MBSFN-based V2V communication

When active road safety support is implemented in C-ITS, common information must be delivered to many vehicles within certain geographic regions; e.g., status information, such as position, velocity and direction, of one vehicle is shared with all other vehicles in the vicinity via CAMs to enhance mutual awareness of road traffic participants. Prior studies have shown that both a direct exchange of periodically generated CAMs using DSRC as well as an indirect distribution over dedicated roadside infrastructure, requires a substantial amount of bandwidth to guarantee timely packet delivery. In fact, with a growing number of vehicles within the geographic region of interest, the currently foreseen 5.9-GHz band for road safety critical communication can easily be overloaded by CAM distribution [9].

In such situations, mobile cellular networks can be helpful to offload some traffic from a dedicated ITS infrastructure. Employing eMBMS/PWS features of LTE, the broadcast nature of cellular systems can be utilized to efficiently deliver the same message to many users within certain geographic regions formed by MBSFN areas. To demonstrate the capabilities of such an approach, we conduct system-level simulations, comparing message distribution via unicast and multicast transmission [10]. Both approaches have certain advantages/disadvantages:

- Unicasting in LTE enables per-user scheduling on favorable time/frequency resources to utilize channel and multiuser

diversity. Furthermore, it supports dynamic link adaptation (multiantenna precoding and transmission rate), as well as selective PHY repetition of lost packets using hybrid automatic repeat request (HARQ). These features facilitate efficient and reliable data transmission to single users with low retransmission latency of lost packets.

- Multicasting in LTE is restricted to employ certain reserved subframes (multicast subframes) and, hence, cannot freely exploit channel diversity. In general, channel and multiuser diversity in multicasting are more difficult to harvest since many users are served in parallel, and, therefore, chances of finding time/frequency resources that are favorable for all users simultaneously are small. Nevertheless, scheduling/resource allocation gains are possible by avoiding the worst time/frequency resources of all users, such that signal outages become less likely. Multicast transmission in LTE does not support PHY retransmission of lost packets and, thus, has to rely on slow (in terms of latency) higher-layer protocols. Furthermore, transmission rate adaptation must be performed according to the channel quality of the worst users to guarantee reliable delivery to all multicast users, reducing the spectral efficiency of the system; indeed, currently existing proposals do not utilize rate adaptation at all but rather rely on fixed rate transmission to avoid CSI feedback from the users. Finally, multicasting in LTE also does not support multiantenna transmission, even though multicast-specific beamforming/precoding and space-time coding has the potential to substantially improve efficiency and reliability of data transmission.

In the following, we investigate the performance of CAM distribution in cellular networks serving vehicles that move at $v = 150$ km/h. We consider fixed-rate transmission (employing CQIs 3, 6, and 9), since rate adaptation in multicasting is not yet supported by the LTE standard. We conduct system-level simulations for an MBSFN area consisting of three macro base stations embedded within a larger cellular network. Within the MBSFN area, CAMs of size 300 bytes are exchanged among 21 vehicles (seven per base station) with a periodicity of 100 ms. Hence, for multicast transmission within the MBSFN area, the total traffic load is $300 \cdot 8 \text{ bit/user} \cdot 21 \text{ users}/100 \text{ ms} \approx 0.5 \text{ Mbit/s}$, since each packet is synchronously broadcast from all three base stations. In case of unicast transmission, on the other hand, each vehicle individually receives the CAMs generated by the other 20 vehicles. Since seven vehicles are attached to each base station of the MBSFN area, this implies a unicast traffic load of $300 \cdot 8 \text{ bit/user} \cdot 20 \text{ users} \cdot 7/100 \text{ ms} \approx 3.4 \text{ Mbit/s}$. We determine the overhead for the cellular network caused by CAM distribution, as well as the operationally critical parameters latency and message loss probability for supporting active road safety in C-ITS. We assume transmission with extended cyclic prefix for unicast and multicast operation.

The results of the simulation are summarized in Table 1. We observe that a unicast transmission with CQI 3 is unable

to support the generated data traffic, i.e., the network overhead caused by CAM distribution is equal to 270% of the capacity (assuming 5-MHz bandwidth); thus, the message loss probability is very high ($> 63\%$) since many CAMs have to be dropped. With multicasting, however, the overhead is reduced to 60% and the message loss probability is below 1%. Yet, even with multicasting in our simulation scenario, it is not possible to sustain the generated CAM network load with CQI less than three, because the transmission efficiency would be too low. Notice that, at most, six subframes per radio frame (consisting of ten subframes) can be reserved for MBSFN operation in LTE. A very important metric for active road safety is latency. Since eMBMS does not support retransmission of lost packets, latency accumulates in multiples of the message generation period (100 ms) in case of packet loss. Thus, even short signal outages can severely increase latency; correspondingly, latency of multicasting deteriorates with increasing transmission rate. For active road safety, latency below 100 ms has to be achieved [11]. Notice that Table 1 only presents radio-link latency for downlink transmissions, neither accounting for uplink from vehicles to base stations nor for CAM distribution within the MBSFN area; hence, we present the percentage of users with downlink latency below 50 ms to incorporate a safety margin with respect to the prescribed 100 ms. We observe in Table 1 that LTE is not able to sustain this latency with sufficiently high probability in our simulation. Hence, performance improvements of multicasting in LTE are required to enable dependable support of road safety applications.

In the section “Multicast Enhancements,” we discuss potential enhancements of LTE to improve multicast transmission, employing dynamic link adaptation and coordination of multipoint transmission. In addition to enhancing the wireless transmission, we also see potential in optimizing the C-ITS protocol itself. As mentioned previously,

Table 1. A comparison of unicast and multicast transmission for CAM distribution in cellular networks.

Rate	Metric	Unicast	Multicast
CQI3 (0.377 bit/sym)	Overhead	270%	60%
	Message loss prob.	$> 63\%$	$< 1\%$
	Latency < 50 ms	n/a	93%
	Latency < 250 ms	n/a	98%
CQI6 (1.176 bit/sym)	Overhead	86%	20%
	Message loss prob.	7%	3%
	Latency < 50 ms	9.5%	80%
	Latency < 250 ms	85%	93%
CQI9 (2.406 bit/sym)	Overhead	42%	10%
	Message loss prob.	11%	11%
	Latency < 50 ms	49%	54%
	Latency < 250 ms	74%	76%

cellular networks can easily be overloaded by CAM distribution, especially when robust MCSs with low spectral efficiency are employed. In such overload situations, reliable distribution of C-ITS messages cannot be guaranteed and, thus, they must be avoided. One possible method could be to adapt CAM generation periodicity according to user mobility; i.e., at low mobility, less frequent vehicle status updates are required than at high mobility. Also, the spatial area in which CAMs are distributed, i.e., the number of base stations involved in CAM exchange, should be carefully chosen and potentially adapted according to traffic load.

Future enhancements and challenges

In this section, we present promising improvements for wireless vehicular communications in cellular networks and highlight associated signal processing challenges. Our focus is on two topics: in the section “CSI Feedback Enhancement,” we discuss CSI feedback enhancements to improve throughput performance of mobile users, employing channel predictive approaches to partly compensate for feedback delay and temporal channel variation. As we demonstrated in the section “Dual Connectivity-Enhanced V2I Transmission,” inaccurate and outdated CSIT can be the limiting factor for the achievable efficiency and reliability of wireless data transmission. Hence, enhancing CSI feedback for high-mobility situations is an essential prerequisite of wireless vehicular communications. In the section “Multicast Enhancements,” we consider techniques for efficient multicast transmission, putting scope not only on advanced multiple-input, multiple-output (MIMO) and coordinated multipoint (CoMP) transmission schemes, but also highlighting issues related to scheduling and resource assignment. Such techniques are important for vehicular scenarios to avoid the network overload problem discussed in the section “MBSFN-Based V2V Communication” by improving the capacity of multicast transmissions. The presented methods, again, require accurate CSIT for beamformer/precoder calculation and, hence, rely on efficient CSI feedback techniques. Finally, we present further important topics for progressing wireless vehicular communications in the section “Further Research Topics to Enhance Wireless Vehicular Communications.”

CSI feedback enhancement

CSI is useful for achieving the highest performance in multiple antenna wireless communications by enabling transmission rate adaptation, adaptive MIMO beamforming, and spatial multiplexing. While CSI at the receiver (CSIR) is comparatively easy to obtain through pilot-aided channel estimation, CSIT in frequency-division duplex systems is only available if the receiver provides feedback information to the transmitter; Figure 4 illustrates the situation. (In time-division duplex systems, channel reciprocity can be exploited to estimate the CSI on both sides of the link.) Since the feedback link between the transmitter and receiver is of limited capacity, quantization of CSI is necessary to enable signaling with a finite number of bits; this is known as *limited feedback operation*. In general, we distinguish between providing explicit and implicit CSI feedback. In the former case, the channel, as estimated by the receiver, is directly quantized and fed back, while in the latter case derived information, such as the optimal MIMO precoder and the supported transmission rate, is provided as side information. In LTE, the second type of feedback information is currently supported with the standard-defined CQI, the precoding matrix indicator (PMI), and the rank indicator (RI). More advanced multiuser MIMO and CoMP schemes, however, mostly rely on explicit CSIT.

Implicit CSI feedback

In the section “Dual Connectivity-Enhanced V2I Transmission,” we observed that CQI feedback, as employed for transmission rate adaptation in LTE, is sensitive to feedback delay in high-mobility scenarios, since outdated CQI feedback can cause signal outages, thereby reducing transmission efficiency and increasing latency. A similar behavior exists for the MIMO-specific feedback information PMI and RI, yet with reduced sensitivity since their mismatch does not directly cause signal outages.

To reduce the impact of feedback delay, either the transmitter or the receiver should attempt to compensate for it. At the receiver side, one natural approach is to apply channel prediction. The simplest method for OFDM-based LTE is a subcarrier-wise linear extrapolation of the channel transfer function between pairs of transmit and receive antennas. If channel statistics are known, the prediction performance can be improved by linear minimum mean squared

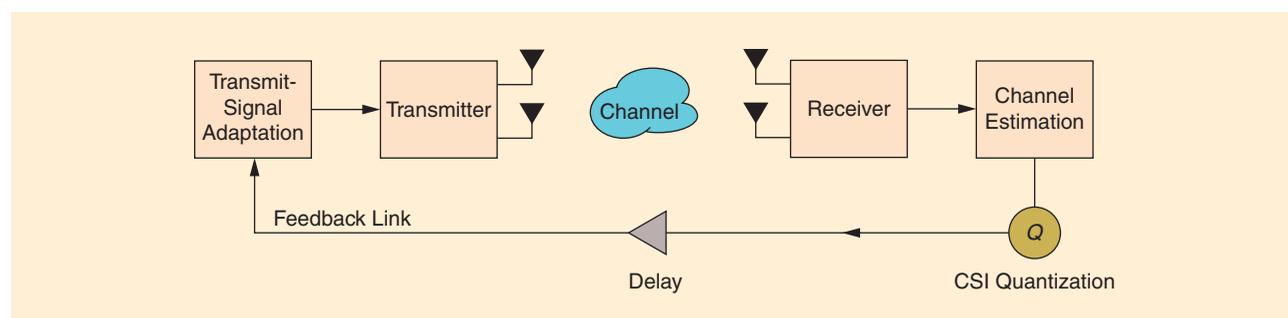


FIGURE 4. An illustration of limited feedback operation in wireless communications assuming error-free but delayed feedback.

error (MMSE) filtering. Without explicit knowledge of the channel statistics, iterative algorithms, such as least mean squares (LMS) or recursive least squares (RLS), can be applied to estimate the MMSE filter coefficients during operation. The linear MMSE filter can further be extended to jointly estimate the channel between all transmit and receive antenna pairs to exploit potential correlations [12]; similarly, the correlation between subcarriers can be utilized by joint estimation of the filters applied on multiple subcarriers.

Alternatively, instead of predicting the channel transfer function, one can predict the channel impulse response and transform the predicted impulse response to the frequency domain for feedback calculation. Efficient prediction of the channel impulse response is possible by applying the minimum-energy band limited discrete prolate spheroidal Slepian sequences as basis functions [13], which are good representatives for the finite support delay-Doppler spreading function of the wireless channel.

If channel prediction at the receiver is too costly, the transmitter can alternatively process the CQI feedback from the user before rate adaptation. One possibility for CQI selection at the transmitter is to maximize the expected throughput at time instant k

$$\text{CQI}[k] = \underset{\text{CQI}_i, i \in \{1, \dots, C\}}{\text{argmax}} T(\text{CQI}_i) \mathbb{P}\{T(\text{CQI}_i) \leq r[k]\}, \quad (2)$$

with $T(\text{CQI}_i)$ denoting the transmission rate associated with CQI_i and C being the number of available CQIs as specified by the standard. The term $\mathbb{P}\{T(\text{CQI}_i) \leq r[k]\}$ is the probability that the currently supported rate $r[k]$ of the channel sustains the selected rate. This probability is commonly not known a priori by the transmitter; yet, assuming that the channel statistics do not vary too quickly, it can easily be learned over time from the user CQI feedback, which signals the highest instantaneously supported MCS [7]. Such a rate adaptation approach can potentially cause unacceptably high block-error ratio and signal outage probability. This can be avoided by adding an additional constraint to problem (2) to account for the acceptable block-error ratio.

In Figure 5, we compare the performance of the described methods in dependence of the normalized feedback delay. We observe that the predictive schemes achieve close to optimal performance up to a certain critical delay, whose exact value depends on the sophistication of the prediction method. Compensation at the transmitter, employing (2) with probabilities estimated from the user feedback, does not achieve optimal performance; yet, it exhibits more robust behavior at very high normalized feedback delay. A further improvement of predictive schemes can be expected by underlying realistic parametric radio channel models to reduce the number of free parameters to estimate/predict [14].

If channel prediction at the receiver is too costly, the transmitter can alternatively process the CQI feedback from the user before rate adaptation.

Explicit CSI feedback

Advanced multiuser MIMO and CoMP transmission schemes, such as block diagonalization (BD) or regularized BD (RBD) precoding, interference-leakage-based precoding [15], and interference alignment (IA), require explicit CSIT for the calculation of precoders. These pre-

coding techniques can enhance the network capacity and reliability by reducing interference among subscribers and base stations. For many such precoding techniques, the underlying CSI can efficiently be represented as a point on a topological manifold. Let $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ denote the channel matrix containing the complex-valued channel gains between the N_t transmit antennas and the N_r receive antennas; we assume $N_t \geq N_r$, which is commonly fulfilled in cellular downlink transmission. For channel-subspace-based precoding techniques, such as BD precoding and IA, the N_r -dimensional subspace spanned by $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is required as CSIT for precoder calculation; this information can be represented as a point on the complex Grassmann manifold $\mathcal{G}(N_t, N_r)$ of N_r -dimensional subspaces in the N_t -dimensional Euclidean space [16]. For other precoding schemes, such as RBD and the interference-leakage-based schemes presented in the section “Multicast Enhancements,” the directions and magnitudes of the channel eigenmodes are necessary CSIT; i.e., the eigenvectors and eigenvalues of the channel Gramian $\mathbf{H}^* \mathbf{H}$, with \mathbf{H}^* denoting the conjugate-transpose of \mathbf{H} . The Gramian is a point on the manifold of symmetric positive semidefinite matrices [17]. Furthermore, the matrix of eigenvectors corresponding to the N_r nonzero eigenvalues of $\mathbf{H}^* \mathbf{H}$ can be interpreted as a point on the compact Stiefel manifold $\mathcal{S}_t(N_t, N_r)$ over the complex numbers [18]. Such CSI representation on manifolds enables dimensionality reduction, which allows efficient quantization with reduced feedback overhead as compared to quantization of the entire channel matrix.

To further reduce the feedback overhead, temporal channel correlation can be exploited through low-rate predictive

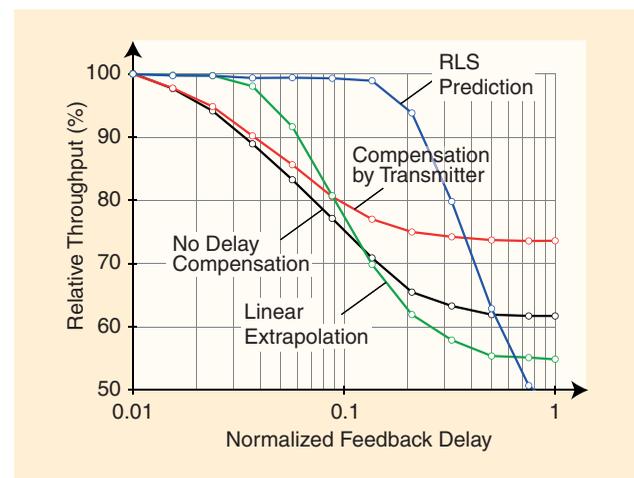


FIGURE 5. The sensitivity of different CQI feedback calculation methods with respect to feedback channel delay and user mobility.

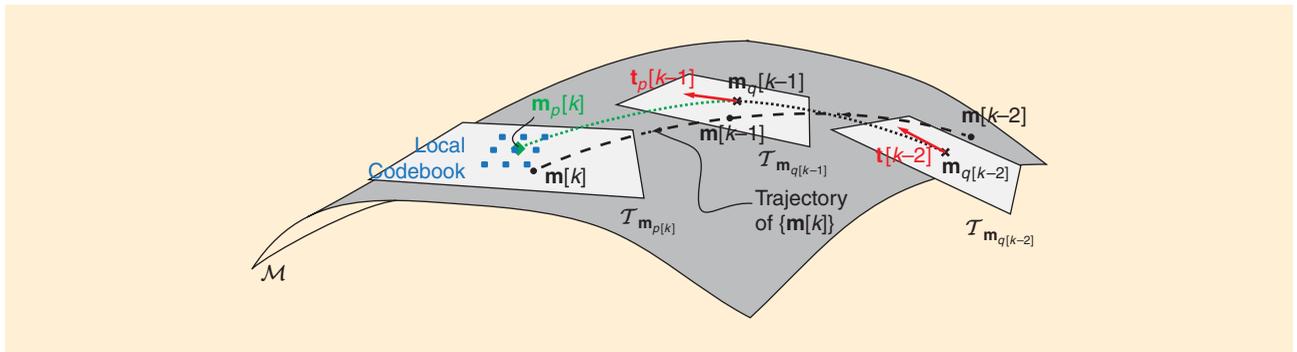


FIGURE 6. An illustration of predictive manifold quantization of a temporally correlated process $\{m[k]\}$ on the manifold \mathcal{M} .

manifold quantization [17]–[19]. Currently, the best rate-distortion performance is achieved by adaptive quantizers that adjust the applied quantization codebook to the temporal evolution of the channel. We summarize this adaptive quantization principle, as introduced for the Grassmannian in [12], for general Riemannian manifolds \mathcal{M} and highlight the most important differential geometric concepts required. For this purpose, we consider a general temporally correlated process of points $\{\mathbf{m}[k]\} \in \mathcal{M}$ that is to be quantized; this process $\{\mathbf{m}[k]\}$ evolves on the manifold \mathcal{M} as illustrated in Figure 6. The basic idea of predictive quantization is to utilize previously quantized observations $\mathbf{m}_q[k-1] = \mathcal{Q}\{\mathbf{m}[k-1]\}$, $\mathbf{m}_q[k-2] = \mathcal{Q}\{\mathbf{m}[k-2]\}$, ... to predict the current point $\mathbf{m}[k]$, where $\mathcal{Q}\{\cdot\}$ denotes the employed quantization function. Given the prediction $\mathbf{m}_p[k]$, the method then generates a local quantization codebook to determine $\mathbf{m}_q[k] = \mathcal{Q}\{\mathbf{m}[k]\}$. Manifold prediction and codebook adaptation are based on quantized CSI, since the decoder at the back end of the feedback link must be able to reproduce these steps. Realizing such a scheme on a manifold requires some additional intermediate actions as detailed next.

First, we notice that a linear prediction, as in the section “Implicit CSI Feedback,” is not meaningful on a general manifold, since an addition of points on the manifold or scalar multiplication with values from an underlying field, as we have on a linear vector space, is in general not defined. Yet, to each point $m \in \mathcal{M}$, an entire linear vector space, the tangent space \mathcal{T}_m , is associated, which locally represents the geometry of the manifold in the Euclidean space. Tangent vectors $\mathbf{t} \in \mathcal{T}_m$ are induced by curves $\gamma(t) \in \mathcal{M}$, such that,

$$\mathbf{t} = \left. \frac{\partial}{\partial t} \gamma(t) \right|_{t=0}, \quad \gamma(0) = \mathbf{m} \quad (3)$$

with scalar $t \in \mathbb{R}$ parameterizing the curve; we assume without loss of generality that the curve traverses \mathbf{m} at $t = 0$. In general, multiple curves can induce the same tangent, since it only depends on the first-order derivative.

For our manifold predictor, we need a one-to-one relationship between pairs of points $\mathbf{m}_1, \mathbf{m}_2$ on the manifold and tangent vectors $\mathbf{t} \in \mathcal{T}_{\mathbf{m}_1}$, such that,

$$\mathbf{t} = \lambda(\mathbf{m}_1, \mathbf{m}_2) \in \mathcal{T}_{\mathbf{m}_1}, \quad \mathbf{m}_2 = \rho(\mathbf{m}_1, \mathbf{t}) \in \mathcal{M}, \quad (4)$$

$$\mathbf{m}_2 = \rho(\mathbf{m}_1, \lambda(\mathbf{m}_1, \mathbf{m}_2)), \quad \forall \mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M}. \quad (5)$$

This is known as a *compatible lifting-retraction pair*, with $\lambda(\cdot, \cdot)$ being the lifting map and $\rho(\cdot, \cdot)$ the retraction map [18]. Ideally, we favor the exponential and logarithmic maps associated with the geodesic curve between $\mathbf{m}_1, \mathbf{m}_2$ for this purpose; yet, these are difficult to evaluate for certain cases, such as the Stiefel manifold. The geodesic represents the shortest path between $\mathbf{m}_1, \mathbf{m}_2$ on the Riemannian manifold.

Given a lifting-retraction pair, we can perform prediction in the linear tangent space, where we can reuse well-known linear algorithms, such as linear prediction and MMSE filtering, and translate the tangent prediction onto the manifold using the retraction. That is, given a set of previous observations $\{\mathbf{m}_q[k-1], \dots, \mathbf{m}_q[k-L-1]\} \in \mathcal{M}$, we calculate

$$\mathbf{t}_i = \lambda(\mathbf{m}_q[k-1], \mathbf{m}_q[k-1-i]) \in \mathcal{T}_{\mathbf{m}_q[k-1]}, \quad \forall i \in \{1, \dots, L\}, \quad (6)$$

$$\mathbf{t}_p[k-1] = \mathcal{P}(\mathbf{t}_1, \dots, \mathbf{t}_L) \in \mathcal{T}_{\mathbf{m}_q[k-1]}, \quad (7)$$

$$\mathbf{m}_p[k] = \rho(\mathbf{m}_q[k-1], \mathbf{t}_p[k-1]) \in \mathcal{M}, \quad (8)$$

with $\mathcal{P}(\cdot)$ denoting the applied prediction function. Finally, we need to construct a local quantization codebook around $\mathbf{m}_p[k]$ for the quantization of $\mathbf{m}[k]$. This can again be achieved most easily by generating a codebook for the Euclidean tangent space and translating it onto the manifold; in [16] and [18], we provide corresponding codebook constructions for the Grassmannian and the Stiefel manifold, respectively.

In Figure 7(a), we demonstrate the performance of such predictive quantization on the Stiefel and Grassmann manifolds in dependence of the normalized sampling interval $\tau_s = t_s/d$, with t_s being the time in between two samples of the manifold process $\{\mathbf{m}[k]\}$. Similar to the normalized feedback delay introduced previously, the normalized sampling interval accounts for the temporal variation of consecutive points of the manifold process $\{\mathbf{m}[k]\}$ due to movement. We assume $N_i \times N_r = 6 \times 2$ and employ a quantization codebook of size 256; i.e., the quantization resolution is

8 bits/sample. The performance of the quantization scheme is gauged in terms of the signal distortion introduced by the quantization process. We employ the chordal distance on the corresponding manifold as distortion metric; see [18] for details. In Figure 7(a), we compare the quantization distortion of predictive quantization to memoryless and differential quantization. Memoryless quantization refers to a quantization scheme that considers each sample $\mathbf{m}[k]$ individually without accounting for the past of the process $\{\mathbf{m}[k]\}$. Differential quantization refers to the case $\mathbf{m}_p[k] = \mathbf{m}_q[k-1]$.

Both differential and predictive quantization achieve an improvement over memoryless quantization for $\tau_s \leq 0.1$; at large τ_s , the distortion of differential/predictive quantization saturates at the performance of memoryless quantization. Hence, the considered differential/predictive quantization schemes achieve a significant gain only at low mobility; e.g., assuming $t_s = 1$ ms, $f_c = 800$ MHz, the point $\tau_s = 0.01$ corresponds to $v = 13.5$ km/h. Since CSI feedback in LTE is currently at most foreseen once every subframe (once every millisecond), this implies that currently available differential/predictive manifold quantization schemes are only advantageous at low mobility and not in vehicular scenarios. Performance can be improved by extending the codebook size; however, complexity issues will ultimately put limits on the supported size.

In Figure 7(b), we show the throughput of BD and RBD precoding when applying predictive Grassmannian and Stiefel manifold quantization to obtain CSIT. We show the performance relative to RBD with perfect CSIT. Imperfect CSIT

Multicasting plays an important role in vehicular communications, since it enables efficient sharing of common information, such as traffic management advices and vehicle status updates, among vehicles.

due to quantization distortion causes residual multiuser interference, degrading the achievable throughput. We observe a significantly better performance of RBD at low signal-to-noise ratio (SNR), since this method implicitly reduces the number of served users with decreasing SNR via power allocation; BD precoding, on the other hand, always serves three users in parallel each with two spatial streams. At high SNR, BD performs slightly better due to the lower quantization distortion of its Grassmannian

quantizer, causing less residual multiuser interference as compared to RBD with Stiefel manifold quantization. Notice that, in this simulation, we assume CSI feedback every millisecond and negligible feedback delay between the users and the base station. In the considered low-mobility scenario with $f_d = 10$ Hz, this is, however, not a significant restriction, because a reasonable feedback delay can be compensated via channel prediction at the receiver; see [12] for a breakdown of the contributions of quantization and delay compensation onto the overall distortion.

At high mobility, the presented CSI feedback method fails to achieve significant multiuser MIMO gains due to insufficient accuracy of CSIT. To support high-mobility scenarios, it would be necessary to reduce the sampling and feedback interval τ_s at least by a factor of ten. Additional gains might also be possible by improving the prediction function in (7); in our simulations, we employed linear prediction in combination with LMS to optimize filter coefficients. Alternatively, precoding schemes that are robust with respect to outdated CSIT, such as retrospective interference alignment, can be utilized. In general, though, such schemes require coding over many

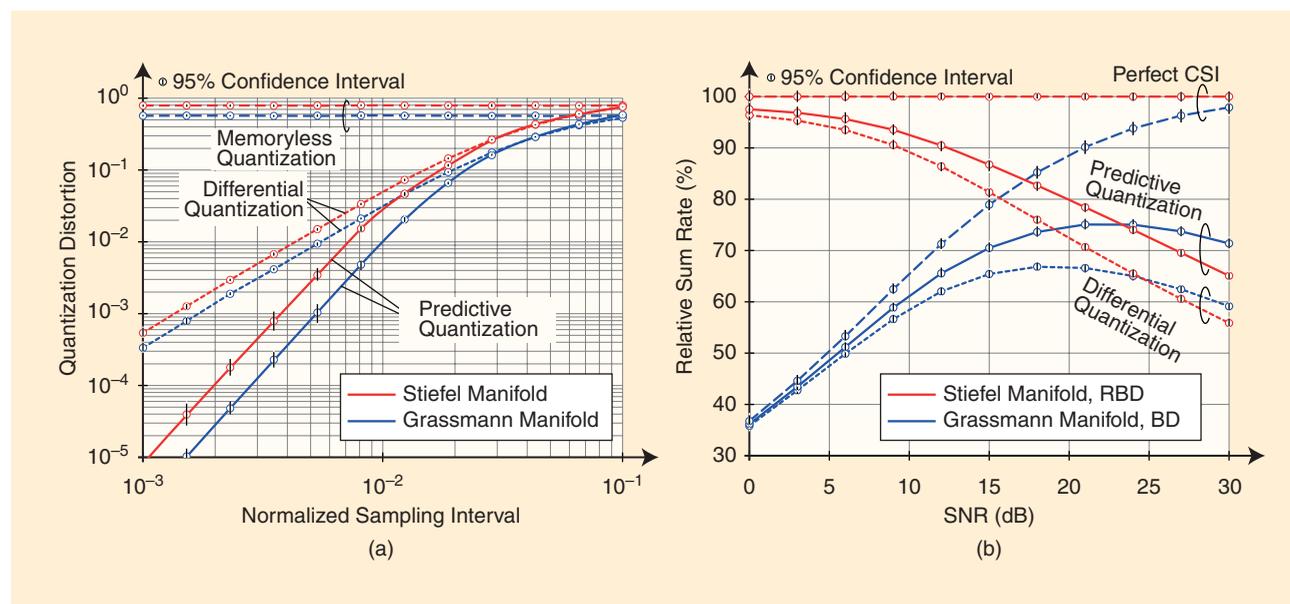


FIGURE 7. The quantization error of (a) manifold quantizers and (b) throughput comparison of RBD and BD with limited feedback at normalized sampling interval $\tau_s = 0.01$ and 8-b quantization codebook. We consider antenna arrays of size $N_t \times N_r = 6 \times 2$.

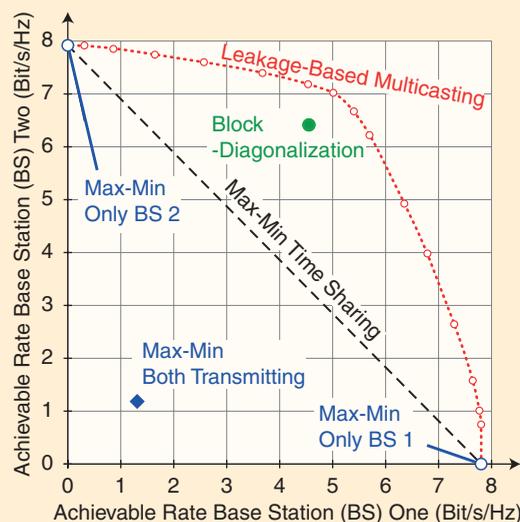


FIGURE 8. The achievable rate-region of the multicast interference channel with two base stations each multicasting common information to six users over $N_t = 8$ transmit antennas. Without beamforming, the SINR is 0 dB.

time frames for efficient operation, which may not be tolerable in terms of latency.

Multicast enhancements

Multicasting plays an important role in vehicular communications, since it enables efficient sharing of common information, such as traffic management advices and vehicle status updates, among vehicles. As we have seen in the section “MBSFN-Based V2V Communication,” though, multicasting of CAMs can easily overload cellular networks, leading to unreliable operation of C-ITS. Next, we discuss improvements of the LTE standard to enhance the efficiency of cellular multicasting. We first consider multicasting-specific MIMO beamforming techniques, which will especially be relevant in future 5G cellular networks employing large-scale full-dimension (FD) MIMO arrays. We then discuss scheduling and resource allocation issues associated to multicast users in LTE networks.

MIMO and coordination

Currently, data transmission in LTE’s eMBMS does not support multiantenna techniques. Already, early investigations on MIMO multicast-beamforming in single-cell networks, however, demonstrate promising capacity gains through max-min beamforming [20]. This approach attempts to maximize the minimum SNR of the multicast users without accounting for interference caused to other users. Even this seemingly simple problem, however, turns out to be nonconvex, and requires semidefinite relaxation (SDR) to determine, in general, suboptimal solutions; only for a small number of transmit antennas and users SDR provides a globally optimal beamforming solution.

Max-min beamforming becomes even more intricate when interference between multiple transmitters is considered, with each of them transmitting information to a different set of multicast users. This scenario is known as the *multicast interference channel*. In cellular networks, it represents a situation where several neighboring base stations serve disjoint sets of multicast users; in vehicular environments, it can come up along highways, where consecutive stretches of the road are served by multiple transmitters. In [21], an interference-leakage-based approach for coordinated beamformer optimization in the multiple-input, single-output multicast interference channel is proposed, where only transmitters are equipped with multiple antennas and not receivers. The method applies max-min beamforming optimization at each base station individually, while restricting the interference leakage caused to users of other base stations. Additionally, it employs an iterative exchange of interference-leakage parameters among base stations in combination with dual gradient optimization over leakage parameters to determine a locally optimal operating point of the multicast interference network. Figure 8 demonstrates the performance of this approach in a multicast interference network with two transmitters each serving six users over $N_t = 8$ transmit antennas. We consider an SNR of 20 dB and assume that, without beamforming, all users receive both transmitters, on average, equally strong; i.e., the macroscopic channel gain with respect to both transmitters is equal. This corresponds to a cell-edge situation where we can expect the largest gains from transmitter coordination. Notice, though, that the method itself is not restricted to equal gain scenarios. Figure 8 shows that, compared to max-min beamforming without considering interference, the rate region can substantially be extended. Hence, the method can significantly improve the cell-edge performance, thereby reducing connectivity failures due to extensive intercell interference. In general, we cannot claim that the method achieves the maximal rate region, since it is only guaranteed to converge to a local optimum; for the shown result, however, we have confirmed global optimality through monotonic optimization utilizing the outer-polyblock algorithm. In the considered high-SNR situation, BD precoding also performs reasonably well. If users are uniformly distributed in the network, the gain over time sharing diminishes with a growing number of users, since the optimal beam pattern tends to be isotropic. Yet, in practical situations, such as vehicles on a road, users are often spatially clustered, which may help mitigating such performance degradation. Early results to support multiple antennas at transmitters and receivers are reported in [22], leading to an iterative alternating optimization of transmit and receive filters. Extensions from beamforming to multiple data stream transmission are yet to be developed.

Beamformer designs for joint multicasting within MBSFN areas formed by multiple base stations, which account for interarea-interference in between neighboring MBSFN areas, are not available so far. Compared to the leakage-based multicasting scheme of [21], this implies additional per base station power constraints for the joint MBSFN beamformer

$$\max_{\mathbf{f} \in \mathbb{C}^{N_{\text{tot}} \times 1}} \min_{u \in U_{\text{MBSFN}}} \text{SINR}_u \quad (9)$$

subject to:

$$\begin{aligned} \|\mathbf{S}_b \mathbf{f}\|_2^2 &\leq P_b, \forall b \in B_{\text{MBSFN}} \\ \text{Leak}_{\bar{u}} &\leq L_{\bar{u}}, \forall \bar{u} \in \bar{U}_{\text{MBSFN}}, \end{aligned}$$

with f denoting the joint beamformer applied over all $N_{\text{tot}} = \sum_{b \in B_{\text{MBSFN}}} N_{i,b}$ transmit antennas $N_{i,b}$ of the set of B_{MBSFN} base stations within the MBSFN area. The parameter SINR_u represents the SINR of user u from the set U_{MBSFN} of users served via multicasting in the MBSFN area and $\text{Leak}_{\bar{u}}$ denotes the interference leakage caused to user \bar{u} from the set \bar{U}_{MBSFN} of users outside the MBSFN area; it is upper bounded by the leakage constraint $L_{\bar{u}}$. Finally, \mathbf{S}_b represents a selector matrix, which chooses the entries of \mathbf{f} corresponding to the transmit antennas of base station b , and P_b is the power constraint of base station b . This problem is nonconvex but, similar to [20], is amenable to SDR. The operating point of this multicast interference network with spatially distributed multicast transmitters can be optimized by coordinating the interference-leakage constraints of interfering base stations. Such coordination can be achieved by applying the dual gradient optimization proposed in [21].

Notice that the discussed beamforming approaches are all based on explicit CSIT, and, hence, the CSI quantization problematic discussed in the section “CSI Feedback Enhancement” applies to these methods as well. Furthermore, since the methods are iterative in nature, they exhibit intrinsic delay to achieve convergence. In vehicular environments, it might be necessary to employ suboptimal heuristics that are derived from optimal schemes; such heuristics are also provided in [21]. Also, especially for road safety applications, outage probability and latency minimization might be of higher importance than rate maximization. Optimization problems for such targets, however, are hard to formulate due to potential nonstationarity of wireless channels in vehicular environments [23]. Still, for limited time frames, assuming quasi-stationary conditions may be valid to formulate outage-constrained optimization problems.

Scheduling and resource allocation

A major weakness of eMBMS in the LTE standard is its inflexibility of multicast transmission; i.e., a fixed amount of subframes within each radio frame is reserved for multicasting. One reason for this is that multicast transmissions within MBSFN areas suffer from a larger delay spread due to single-frequency transmission of the same signal from multiple spatially distributed base stations. To compensate for the increased delay spread, multicast subframes employ the extended cyclic prefix (16.7 μs) of LTE, whereas unicast subframes mostly use the shorter normal cyclic prefix (4.7 μs) to minimize waste of bandwidth. According to our work on link adaptation, however, utilizing the extended cyclic prefix may not pay off, even if intersymbol interference (ISI) occurs; in many cases of realistic SNR, it is more efficient to simply

accept the remaining ISI and use a more robust transmission rate to compensate for it [24]. Hence, the LTE standard should be extended to support multicasting in MBSFN areas with a short cyclic prefix. This then also allows the incorporation of multicast transmissions into normal subframes, facilitating a better dynamic resource assignment according to traffic requirements, as well as the exploitation of channel diversity in multicasting. Within the 3GPP, such extensions are considered in the study item on single-cell point-to-multipoint transmission; within this study, single-cell multicasting over the downlink shared channel is evaluated, which until now has only handled unicast data transmissions.

A big challenge from a signal processing perspective is efficient scheduling and coordination of multicast transmissions. In the context of vehicular communications, this involves dynamic formation of MBSFN areas, to limit the amount of CAMs that must be distributed while still providing mutual awareness over sufficiently large geographic areas. Furthermore, it implies determining groups of multicasting users that must share CAMs due to close spatial proximity. The multicast scheduler should also be able to decide to offload users with very poor channel quality to unicast transmissions, to support selective PHY retransmissions via HARQ, such as to reduce latency and to improve reliability. First results in this direction are provided in [25], where the authors consider optimizing proportional fairness of concurrent multicast and unicast transmissions. Heterogeneity of user channel conditions is considered in [25] by optimally partitioning multicast users into groups, so that users with good signal strength do not suffer by being grouped together with users of poor signal strength. A further challenge comes up when the system supports direct V2V transmission in addition to cellular-assisted transmission; then the scheduler must further group vehicles according to direct and assisted transmission.

Further research topics to enhance wireless vehicular communications

Channel estimation at high mobility

At very high mobility, accurate CSIR estimation using pilot signals can become challenging due to strong temporal variation of the wireless channel. Prior investigations have shown that LTE does not achieve MIMO spatial multiplexing gains at high mobility due to insufficient density of pilot symbols in the time domain [26]. This low density of pilot symbols can cause poor accuracy of channel estimation and, as a consequence, unreliable symbol detection. A remedy can be provided by adapting the pilot pattern to the Doppler- and delay-spread of the channel. These two properties characterize the coherence time and bandwidth of the channel, respectively, i.e., the time/frequency intervals over which the channel stays approximately constant. Such an adaptive pilot scheme can efficiently be realized with minimal feedback information about coherence time and bandwidth from the users, requiring an update whenever channel statistics vary significantly [26].

In vehicular environments, especially on highways, the wireless channel behaves markedly different as compared to other common situations in mobile communications. More specifically, the propagation is characterized by shadowing through other vehicles, high Doppler shifts with often sparse Doppler spectrum due to few dominant scatterers (e.g., road signs, highway overpasses), and possibly inherent nonstationarity of the channel statistics. Such effects can cause substantial performance degradation of common least squares and MMSE channel estimators if they are not properly considered in the algorithmic design. In [23], the impact of the shape of the Doppler spectrum on the performance of channel estimation at the receiver is discussed in more detail. Similarly, timing and carrier synchronization in vehicular scenarios with few dominant scatterers of similar strength that are strongly delay- and Doppler-shifted with respect to each other can be challenging.

ICI mitigation

OFDM multicarrier modulation, as employed in LTE, suffers at high mobility from intercarrier interference (ICI) due to the Doppler spread of the transmit signal. This effect, however, can, to a large extent, be mitigated through iterative ICI estimation and cancellation at the receivers. In [27], the authors propose ICI mitigation algorithms that enable achieving the performance of interference-free transmission. Alternatively, optimal pulse shaping at the transmitter, as currently pushed by many research groups and companies for 5G mobile communications, can be applied to maximize achievable data rates by trading off residual ICI/ISI for spectral efficiency [28]. Such an approach appears especially interesting when waveform parameters, such as subcarrier spacing, prototype pulse shape, and TTI length, can be adapted to the time/frequency-dispersion characteristics of the channel. A major challenge then is to optimize sets of compatible parameters that enable efficiently serving users with strongly different channel properties in parallel, e.g., static and highly mobile users. Novel multicarrier transmit waveforms, employing filter banks [filter-bank multicarrier modulation (FBMC)] or subband filters (universal filtered multicarrier/OFDM), support the necessary flexibility to adjust waveform parameters over subbands and, thus, provide the basis for channel-adaptive modulation.

Multiconnectivity

To enhance the robustness of the wireless transmission link, so as to support highly reliable communication, macro diversity can be exploited by extending the dual connectivity concept of LTE to a multitude of radio network access points. Maintaining multiple parallel connections to several macro base stations and/or small cells promises increased data throughput over multiple parallel data streams, improved reliability due to a reduced outage probability and enhanced robustness with respect to mobility, since hard handovers can be avoided. Yet, it

A big challenge from a signal processing perspective is efficient scheduling and coordination of multicast transmissions.

implies that single users occupy resources at multiple transmission points, causing a reduction of the overall network capacity. To maintain the efficiency of the network, it is therefore necessary to dynamically decide whether the gain of utilizing multiconnectivity for a user outweighs the overhead caused on network capacity. Hence,

multiconnectivity comes at the cost of requiring a sophisticated coordination of multiple transmission points, potentially implying a substantial backhaul signaling overhead. Such enhanced self-organizing network features can, e.g., be realized by cloud radio access network architectures. In [29], the authors investigate the tradeoff between radio-link failures and user throughput in dependence of the size of the set of active multiconnectivity transmission points, showing that both the cell-edge user throughput as well as the mobility performance can be improved simultaneously.

FD MIMO beamforming for high mobility

FD MIMO refers to wireless transmission systems that support active two-dimensional antenna arrays with a large number of antenna elements. This enables high-resolution adaptive beamforming in both the elevation and the azimuth domain, to achieve space-division multiple access gains through spatial separation of users. Within LTE standardization, work is ongoing to implement FD MIMO within Release 14. Currently, hybrid transceiver architectures are of interest, where part of the signal processing is performed in base band and part in the analog domain to limit the number of required radio-frequency chains. Analog beamforming approaches are mostly based on signal azimuth and elevation arrival/departure angles. To implement such schemes at high mobility, it is necessary to account for uncertainty in the signal arrival/departure angles due to user movement and estimation errors. In [30], the authors propose a corresponding robust beamformer optimization problem and demonstrate improved robustness at high mobility.

Conclusions

Vehicular communications is an integral part of innovative transport telematics systems for traffic management and active road safety. It plays a key role in making public and private transportation faster, more reliable, more efficient, and safer. Realizing the necessary information exchange among roadside infrastructure and vehicles efficiently and reliably can be challenging. In that respect, cellular networks can provide valuable support to dedicated vehicular communication systems, since today's cellular base stations are almost ubiquitously accessible and supply high bandwidth wireless connectivity. In this article, we have surveyed ongoing efforts and developments within the 3GPP to implement vehicular communications over LTE. We have discussed research challenges associated with wireless communications at high mobility, and we have provided an overview of promising signal processing techniques to tackle important hurdles. Even though significant progress in enhancing

wireless vehicular communications has been made in the past, there is room left for improvement.

Acknowledgments

The financial support by the Austrian Federal Ministry of Science, Research, and Economy and the National Foundation for Research, Technology, and Development is gratefully acknowledged. The research has been cofinanced by the Czech Science Foundation, Project Number 17-18675S “Future Tranceiver Techniques for the Society in Motion,” and by the Czech Ministry of Education in the frame of the National Sustainability Program under grant LO1401.

Authors

Stefan Schwarz (sschwarz@nt.tuwien.ac.at) received his Dipl.-Ing. degree in electrical engineering in 2009 from the Vienna University of Technology (TU Wien). He received his Dr. techn. degree in telecommunications engineering in 2013 from TU Wien. He has a postdoctoral position at the Institute of Telecommunications of TU Wien, heading the Christian Doppler Laboratory for Dependable Wireless Connectivity for the Society in Motion.

Tal Philoso (tal.philosof@gm.com) received the M.Sc. and Ph.D. degrees from Tel-Aviv University, Israel, in 2003 and 2010, respectively. During 2004–2009, he served as a research fellow in several academia-industry scientific programs supported by the Israel Ministry of Industry. He is currently with the Wireless Enablers Lab, General-Motors R&D Advanced Technical Center Israel.

Markus Rupp (mrupp@nt.tuwien.ac.at) received his Dipl.-Ing. degree in 1988 from the University of Saarbrücken, Germany, and his Dr.-Ing. degree in 1993 at Technische Universität Darmstadt, Germany. Until 1995, he had a postdoctoral position at the University of Santa Barbara, California. From 1995 to 2001, he was with the Wireless Technology Research Department of Bell-Labs, New Jersey. Since October 2001, he has been a full professor of digital signal processing in mobile communications at TU Wien.

References

- [1] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro, “LTE for vehicular networking: A survey,” *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 148–157, May 2013.
- [2] H. Seo, K. D. Lee, S. Yasukawa, Y. Peng, and P. Sartori, “LTE evolution for vehicle-to-everything services,” *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 22–28, June 2016.
- [3] Y. L. Tseng, “LTE-advanced enhancement for vehicular communication,” *IEEE Wireless Commun.*, vol. 22, no. 6, pp. 4–7, Dec. 2015.
- [4] Y. Zhou, Z. Lei, and S. H. Wong, “Evaluation of mobility performance in 3GPP heterogeneous networks,” in *Proc. IEEE 79th Vehicular Technology Conf.*, May 2014, pp. 1–5.
- [5] S. Jha, K. Sivanesan, R. Vannithamby, and A. Koc, “Dual connectivity in LTE small cell networks,” in *Proc. Globecom Workshops*, Dec. 2014, pp. 1205–1210.
- [6] M. Rupp, S. Schwarz, and M. Taranetz, *The Vienna LTE-Advanced Simulators: Up and Downlink, Link and System Level Simulation*. New York: Springer-Verlag, 2016.
- [7] S. Schwarz and M. Rupp, “Throughput maximizing feedback for MIMO OFDM based wireless communication systems,” in *Proc. Signal Processing Advances Wireless Communications*, San Francisco, CA, June 2011, pp. 316–320.
- [8] T. Rappaport, “Wireless communications: Principles and practice,” in *Prentice Hall Communications Engineering and Emerging Technologies Series*. Upper Saddle River, NJ: Dorling Kindersley, 2009.

- [9] L. Shi and K. W. Sung, “Spectrum requirement for vehicle-to-vehicle communication for traffic safety,” in *Proc. IEEE 79th Vehicular Technology Conf.*, May 2014, pp. 1–5.
- [10] I. Safiulin, S. Schwarz, and M. Rupp, “System level simulation of LTE MBSFN networks with high mobility users,” in *Proc. 22nd Int. Conf. Systems, Signals and Image Processing*, London, 2015, pp. 21–24.
- [11] G. Karagiannis, O. Altintas, E. Ekici, G. Heijenk, B. Jarupan, K. Lin, and T. Weil, “Vehicular networking: A survey and tutorial on requirements, architectures, challenges, standards and solutions,” *IEEE Commun. Surv. Tutor.*, vol. 13, no. 4, pp. 584–616, 2011.
- [12] S. Schwarz, R. Heath, Jr., and M. Rupp, “Adaptive quantization on a Grassmann-manifold for limited feedback beamforming systems,” *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4450–4462, 2013.
- [13] T. Zemen, C. F. Mecklenbrauker, and B. H. Fleury, “Time-variant channel prediction using time-concentrated and band-limited sequences,” in *Proc. IEEE Int. Conf. Communications*, June 2006, vol. 12, pp. 5660–5665.
- [14] A. Heidari, A. K. Khandani, and D. Mcavoy, “Adaptive modelling and long-range prediction of mobile fading channels,” *IET Commun.*, vol. 4, no. 1, pp. 39–50, Jan. 2010.
- [15] J. Park, G. Lee, Y. Sung, and M. Yukawa, “Coordinated beamforming with relaxed zero forcing: The sequential orthogonal projection combining method and rate control,” *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3100–3112, June 2013.
- [16] S. Schwarz, R. Heath, Jr., and M. Rupp, “Adaptive quantization on the Grassmann-manifold for limited feedback multi-user MIMO systems,” in *Proc. 38th Int. Conf. Acoustics, Speech and Signal Processing*, Vancouver, Canada, May 2013, pp. 5021–5025.
- [17] D. Sacristan-Murga and A. Pascual-Iserte, “Differential feedback of MIMO channel Gram matrices based on geodesic curves,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 3714–3727, 2010.
- [18] S. Schwarz and M. Rupp, “Predictive quantization on the Stiefel manifold,” *IEEE Signal Process. Lett.*, vol. 22, no. 2, pp. 234–238, 2015.
- [19] R. Pitaval and O. Tirkkonen, “Joint Grassmann-Stiefel quantization for MIMO product codebooks,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 210–222, 2014.
- [20] N. Sidiropoulos, T. Davidson, and Z. Q. Luo, “Transmit beamforming for physical-layer multicasting,” *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, June 2006.
- [21] S. Schwarz and M. Rupp, “Transmit optimization for the MISO multicast interference channel,” *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4936–4949, 2015.
- [22] S. Schwarz, T. Philoso, and M. Rupp, “Leakage-based multicast transmit beamforming,” in *Proc. Int. Conf. Communications*, London, U.K., June 2015, pp. 2405–2410.
- [23] C. Mecklenbrauker, A. Molisch, J. Karedal, F. Tufvesson, A. Paier, L. Bernado, T. Zemen, O. Klemp, and N. Czink, “Vehicular channel characterization and its implications for wireless system design and performance,” *Proc. IEEE*, vol. 99, no. 7, pp. 1189–1212, July 2011.
- [24] E. Zöchmann, S. Pratschner, S. Schwarz, and M. Rupp, “Limited feedback in OFDM systems for combating ISI/ICI caused by insufficient cyclic prefix length,” in *Proc. 48th Asilomar Conf. Signals, Systems and Computers*, Asilomar, CA, Nov. 2014, pp. 1–5.
- [25] J. Chen, M. Chiang, J. Erman, G. Li, K. K. Ramakrishnan, and R. K. Sinha, “Fair and optimal resource allocation for LTE multicast (eMBMS): Group partitioning and dynamics,” in *Proc. IEEE Conf. Computer Communications*, Apr. 2015, pp. 1266–1274.
- [26] M. Simko, P. Diniz, Q. Wang, and M. Rupp, “Adaptive pilot-symbol patterns for MIMO OFDM systems,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4705–4715, Sept. 2013.
- [27] R. Nissel and M. Rupp, “Doubly-selective MMSE channel estimation and ICI mitigation for OFDM systems,” in *Proc. IEEE Int. Conf. Communications*, London, U.K., June 2015, pp. 6295–6300.
- [28] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, “Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency,” *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 80–93, Nov. 2014.
- [29] F. B. Tesema, A. Awada, I. Viering, M. Simsek, and G. P. Fettweis, “Mobility modeling and performance evaluation of multi-connectivity in 5G intra-frequency networks,” in *Proc. IEEE Globecom Workshops*, Dec. 2015, pp. 1–6.
- [30] S. Schwarz and M. Rupp, “Limited feedback based double-sided full-dimension MIMO for mobile backhauling,” in *Proc. 50th Conf. Signals Systems and Computers*, Asilomar, CA, 2016.
- [31] 3GPP. “Technical specification group services and system aspects; Study on LTE support for V2X services (Release 14). [Online]. Available: <http://www.3gpp.org/DynaReport/22885.htm>
- [32] 3GPP. “Technical specification group radio access network; Study on LTE-based V2X services (Release 14). [Online]. Available: <http://www.3gpp.org/DynaReport/36885.htm>

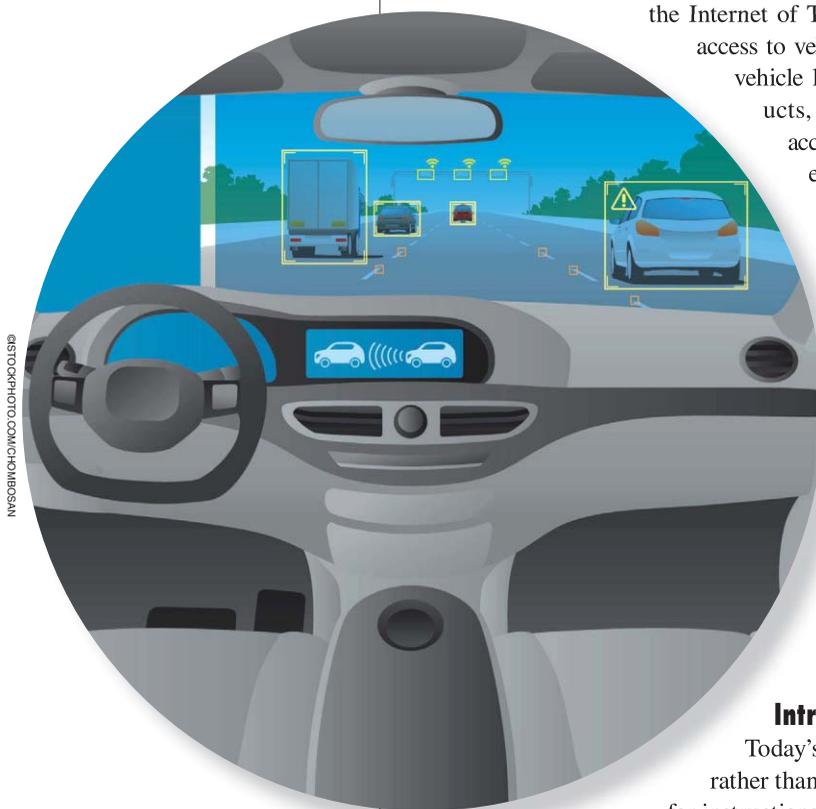
The Future of Automotive Localization Algorithms

Available, reliable, and scalable localization: Anywhere and anytime

Most navigation systems today rely on global navigation satellite systems (GNSS), including in cars. With support from odometry and inertial sensors, this is a sufficiently accurate and robust solution, but there are future demands. Autonomous cars require higher accuracy and integrity. Using the car as a sensor probe for road conditions in cloud-based services also sets other kind of requirements. The concept of the Internet of Things requires stand-alone solutions without access to vehicle data. Our vision is a future with both in-vehicle localization algorithms and after-market products, where the position is computed with high accuracy in GNSS-denied environments. We present a localization approach based on a prior that vehicles spend the most time on the road, with the odometer as the primary input. When wheel speeds are not available, we present an approach solely based on inertial sensors, which also can be used as a speedometer. The map information is included in a Bayesian setting using the particle filter (PF) rather than standard map matching. In extensive experiments, the performance without GNSS is shown to have basically the same quality as utilizing a GNSS sensor. Several topics are treated: virtual measurements, dead reckoning, inertial sensor information, indoor positioning, off-road driving, and multilevel positioning.

Introduction

Today's positioning systems are intended for humans rather than machines. The position is presented and used for instructions in navigation systems or for reporting vehicle data, also including emergency accident location. We refer to the area as *localization algorithms* for several reasons. First, the word *algorithm* indicates software development. Today there is already sufficient information at hand, in terms of sensors and databases, to make a leap in performance compared to GNSS-based solutions. Second, localization is not a system, rather it is a service required by many systems. Third, the term *navigation* is avoided since this is only one application of localization algorithms. Fourth, *localization* is sometimes a more appropriate term



©ISTOCKPHOTO.COM/CHOMBOBAN

than *positioning*, since a true longitude and latitude position is of no value unless the map and situational awareness have the same absolute accuracy.

Consider the schematic picture of a vehicle in Figure 1. The trend is to make vehicles autonomous [1]–[4] and utilize advanced driver assistance systems (ADAS). Hence, there is a need to improve both localization and velocity estimation systems. Basically, it is going beyond traditional point estimation methods [5], [6] to get a better probabilistic understanding [7]–[10] of the environment using more detailed models and filters. The actuators (brake, steering wheel, engine torque) have essentially been the same since the automobile was invented, and only a few new actuator concepts have been introduced (active suspension, movable headlights, etc.).

In stark contrast to the actuators, the number of sensors has increased substantially over the last decade [11]–[13], e.g.,

- the inertial measurement unit (IMU) [14] in the engine control unit (ECU) and in suspension sensors for estimating the vehicle state
- vision, stereo vision, night vision, radar, sonar for monitoring the surroundings, and keeping the vehicle in the lane at a safe distance (i.e., relative position control)
- the wheel speed sensors (WSS) introduced with the anti-lock braking system (ABS) are one of the most versatile sensors in the car
- databases such as vectorized road maps [15]–[18] utilized for positioning including road height, map matching [11], [19]–[22], and pothole indications [23], etc.

Cars are slowly following the development of smartphones. Today there are many radio receivers in vehicles: cellular network, Bluetooth, and Wi-Fi. These can be used in various signal processing applications such as localization and speed estimation. It is less explored that these information sources all include indirect information about the position. The vehicle state sensors contain information of road signatures (curves, banking, slopes, and small variations in the surface height). The vision sensors can see landmarks of known position. How the WSS can be used for positioning is described next.

Sensor fusion is used in all of the aforementioned cases for refining the information, where there are several good examples of virtual (or soft) sensors that compute physical quantities that cannot readily be measured by sensors. Examples include the detection of obstacles, pedestrians, and animals from vision sensors and tire pressure and road friction from WSS. Our approach is based on statistical signal processing techniques, based on a simple odometric model of the vehicle and a model of each sensor relating to the vehicle state. In particular, the road map information is nonlinear and cannot be approximated with a linear Gaussian model, so a PF framework is preferred to Kalman filter (KF) algorithms. The sensor fusion concept is summarized in Figure 2.

Future localization algorithm applications

This section discusses the need for improved localization algorithms highlighting areas such as cloud-based computations, autonomous driving, handheld devices, and mapping.

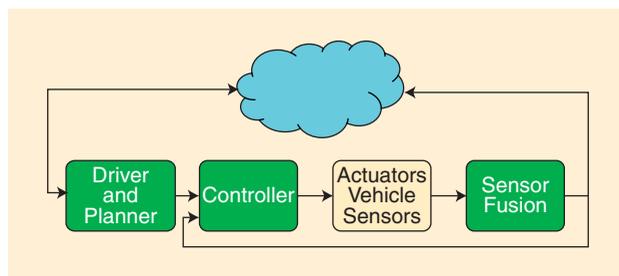


FIGURE 1. An illustration of data flow in a vehicle. Future ADAS functionality might include cloud information as well as control, sensor fusion, and planning.

Cloud-based services

To some extent, positioning today is used for cloud-based crowdsourcing, such as in apps for pothole detection and speed camera positions, among others. This is an area that probably will explode in the future when manufacturers integrate these reports in their own servers and offer their own and other customers services based on this information.

Consider potholes as an example of a virtual sensor: they are annoying to passengers and may be a hazard to the vehicle. These are easily detected by accelerometers, WSS, or suspension sensors, and the presence of potholes can be included in the car's navigation system. The problem is how to share the information between users. Figure 3 shows an illustration of pothole detection and clustering [23]. Many vehicles have, in this case, hit the same potholes and delivered the estimated position to a cloud database. A cloud-based clustering algorithm is then used to merge the various pothole detections into one unique pothole, and possibly also project the position to the road. This information can now be shared with other drivers, but it could also be used by road authorities for maintenance.

Autonomy

Future autonomy will put high demands on the localization algorithms. Despite the media success of self-driving cars, the technology is still in development. On one hand, there is the Defense Advanced Research Projects Agency (DARPA) generation of cars where localization is based on a laser scanner, however, the cost of these vehicles is still far from affordable; see the DARPA grand challenge and urban challenge [1], [2]. Further, the laser scanner's raised placement on the rooftop is not well aligned with the design. Google's self-driving car [3] is equipped with laser, radar, and cameras on the rooftop. Apart from most vehicles, it is not designed to be, or even possible to, drive manually. On the other hand, we have seen self-driving cars (Audi RS7 Piloted Driving presented in Hockenheim, Germany) positioned using differential GPS, including yaw estimation from multiple antennas and camera information. These cars have demonstrated spectacular performance on restricted accurately mapped areas.

Autonomous functions in the car, and in the extreme self-driving cars, will need another level of integrity. The localization algorithms must work in tunnels, parking garages, urban street

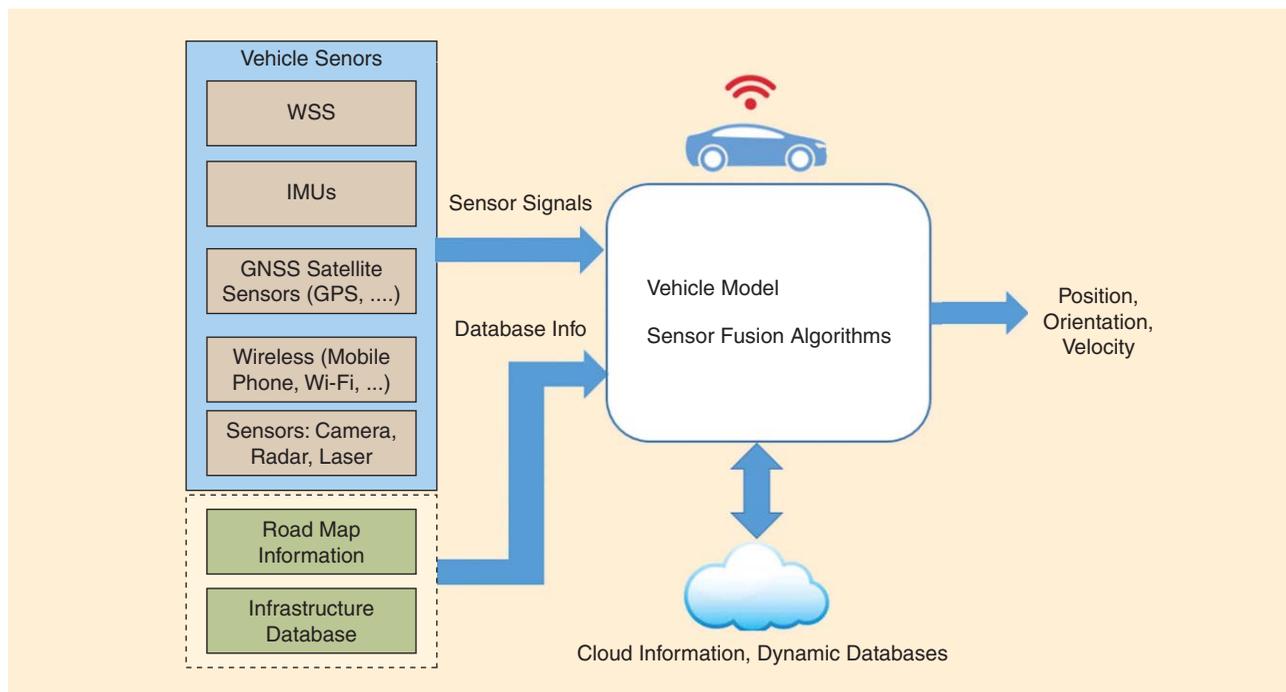


FIGURE 2. An overview of positioning, orientation, and velocity estimation utilizing all available in-vehicle sensors, external databases, and cloud interaction.

canyons, and other areas where GNSS is problematic. If the satellite signal is only mitigated and pseudoranges are available multiple model filters and map constraints might be an option [24]. However, for the general case: indoor driving, long tunnels, and multiple levels, the focus is on map-aided positioning without satellite signals.

Localization must also be robust against jamming and spoofing. Fleet management and theft tracking systems should not rely on access to GNSS.

Devices

With tens of billions of connected devices around us, some of them will be used in vehicles. There is, of course, a demand to keep track of those devices. In some cases, the devices can be connected to the car to take advantage of the sensor information

on the computer area network (CAN) bus. However, there is no standard for the protocol here, so making devices connect to many original equipment manufacturer vehicles is a challenge; i.e., there is also a need for completely stand-alone localization algorithms. It will also be more common to transition from in-vehicle estimation to handheld devices.

Mapping

For positioning/localization, there could either be a priori map information available, or it could be derived from sensor data. Usually, accurate vectorized maps of the road network are available for vehicle positioning. This is the focus on the applications described in this article. Also, many landmarks such as speed-limit signs, etc. could be considered known and available in complementary databases.

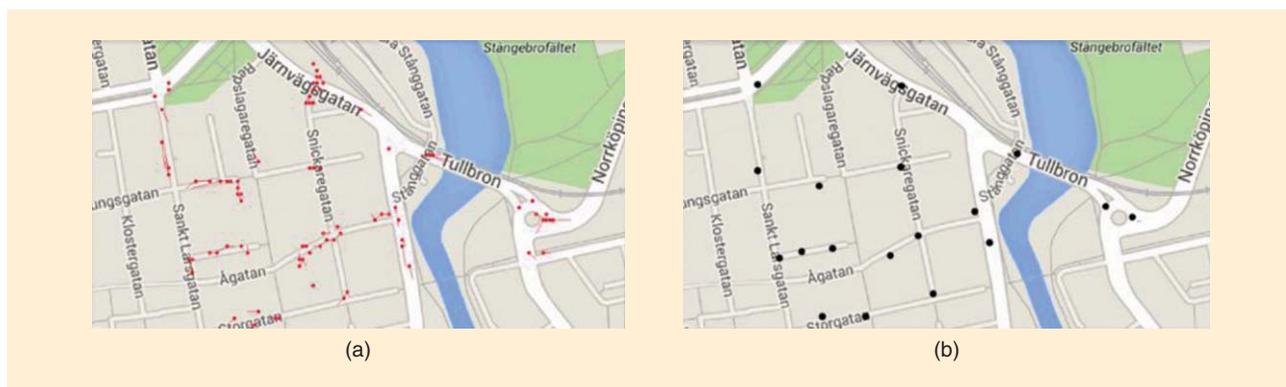


FIGURE 3. There are different sensors in a modern vehicle that can detect potholes. (a) Reports from a fleet of vehicles from NIRA Dynamics are sent to a cloud database [23]. (b) Since the potholes notion of position differs for natural reasons (based on GNSS), clustering is needed in the cloud.

For the sake of completeness, and since the algorithms and methods are closely related, we will briefly discuss simultaneous localization and mapping (slam). It is an extension of the localization problem to the case where the environment is unmodeled and has to be mapped on-line. A survey on the slam problem is given in [25]–[27]. The Fastslam algorithm introduced in [28] has proved to be an enabling technology for such applications. Fastslam can be seen as a special case of Rao–Blackwellized PF (rbpf) or marginalized PF (mpf) (see the section “Bayesian Filtering”), where the map state containing the positions for all landmarks used in the mapping can be interpreted as a linear Gaussian state. The main difference is that the map vector is a constant parameter with a dimension increasing over time, rather than a time-varying state with a dynamic evolution over time. In [29], slam is used to get high accuracy map information (centimeter resolution) utilizing all available sensors such as gps, odometer, and laser. In [30], a different technique utilizing image data for high-accuracy navigation is utilized.

Road maps and map matching

The unique feature with automotive localization algorithms is that vehicles spend most of their time on roads, and this is

also the common theme in this article. We will show how road maps can be integrated with sensor fusion techniques to provide an accurate position with high integrity.

The classical method to improve localization performance is map matching [31], [22]. Here, the position estimate computed from the sensors (for instance, GPS) is mapped to the closest point on the road. This is an appropriate method for presentation purposes, but it suffers from two problems. First, it does not take the topography of the map into account, which implies that the localization can jump from one road to another. Second, the motion dynamics of the vehicle is not combined with the map information in an optimal way. Having said that, there are different types of map matching, basically using the estimated trajectory in combination with the GPS measurement to retrieve the most likely position. They are sometimes referred to as *point-to-point*, *point-to-curve*, and *curve-to-curve matching* [11], [19]–[21]. For a detailed survey over map matching, we refer to [22].

The purpose of this section is to survey different methods, which we refer to as *dynamic map matching*. This includes combining a motion model, sensor models, and the road model in a nonlinear filter, including uncertainties. The problem is fitting a distorted and noisy trajectory to the road network. Figure 4

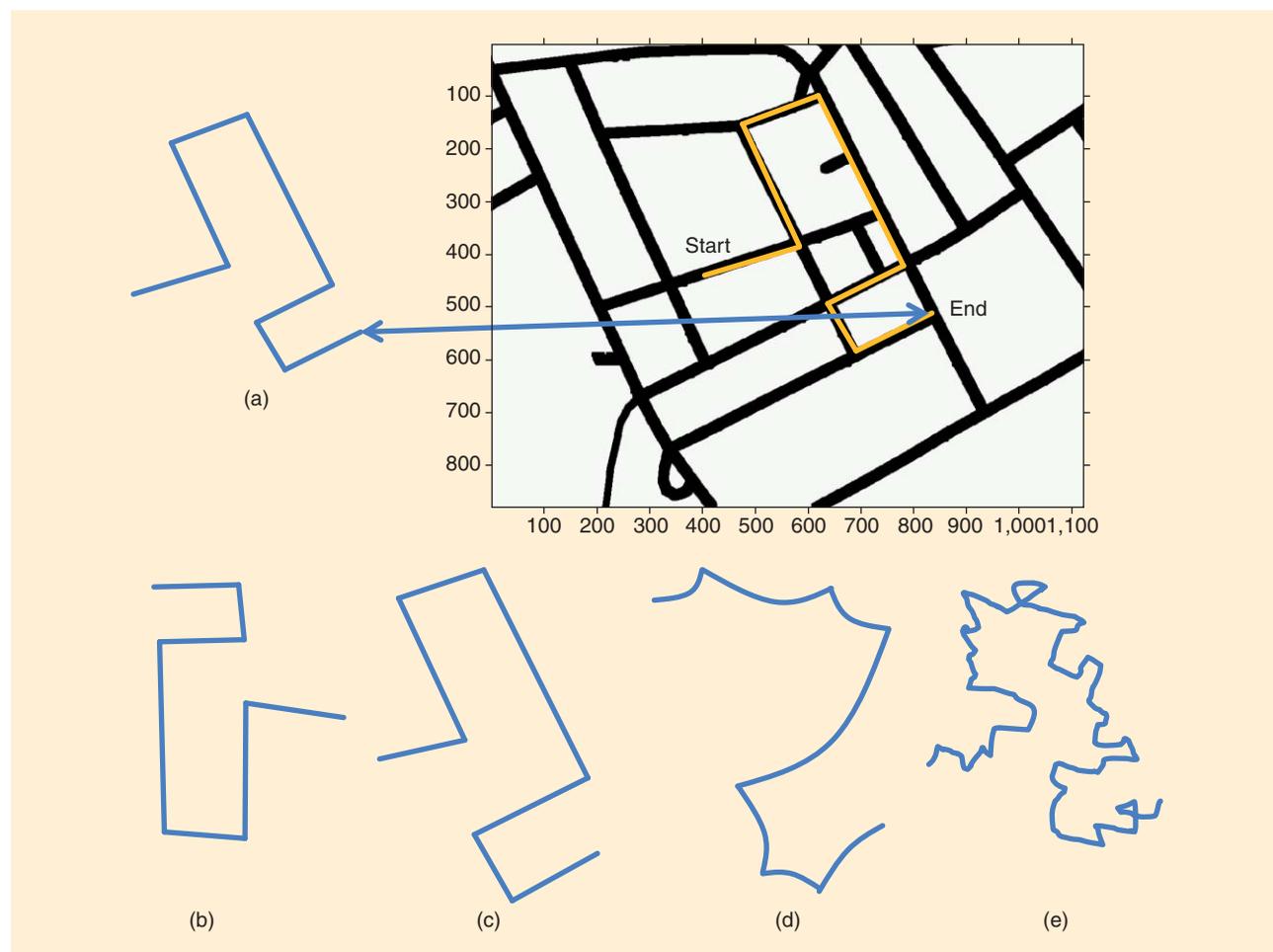


FIGURE 4. The key idea in dynamic map matching is to fit an observed trajectory to the road network. (a) Undistorted trajectory. (b) Undistorted trajectory with random rotation. (c) Trajectory based on biased speed. (d) Trajectory based on biased yaw rate. (e) Trajectory with random noise.

illustrates the principle. Hence, it is possible to utilize only odometry and map information to get an accurate localization [32], [33].

Dead-reckoning principles

Dead reckoning is essentially calculating the integral of velocity or acceleration signals, with or without a vehicle model. It can be based on IMU data or WSS signals, for instance. We will look at several aspects including:

- odometry or dead reckoning based on WSS
- inertial sensor data dead reckoning
- utilization of map matching
- dynamic filtering.

Dead reckoning: Odometry

A simple motion model is based on a state vector consisting of position X , Y , and course ψ , in which case the principle of dead reckoning can be applied. Inserting the observed speed $\vartheta^m(t)$ and angular velocity $\dot{\psi}^m(t)$ as input signals gives the following dynamic model with process noise $w(t)$:

$$\begin{aligned} X(t+T) &= X(t) + \vartheta^m(t)T \cos(\psi(t)) + T \cos(\psi(t))w_\vartheta(t), \\ Y(t+T) &= Y(t) + \vartheta^m(t)T \sin(\psi(t)) + T \sin(\psi(t))w_\delta(t), \\ \psi(t+T) &= \psi(t) + T\dot{\psi}^m(t) + T w_\psi(t). \end{aligned}$$

This model has the following structure ($T = 1$):

$$x_{t+1} = f(x_t, u_t) + g(x_t, u_t)w_t, \quad u_t = (\vartheta_t^m, \dot{\psi}_t^m)^T.$$

Normally, additional sensors are needed to get observability of the absolute position. However, the vectorized road map [15]–[18] contains sufficient information in itself. Note that the speed and the angular velocity measurements are modeled as inputs rather than measurements. This is in accordance with many navigation systems, where inertial measurements are dead reckoned in similar ways. Alternative road graph models are discussed in [34], and second-order motion models in [35].

Dead reckoning: Inertial sensors

Using an IMU, it is possible to directly measure the acceleration and angular rotation. From these measurements it is theoretically possible to integrate the underlying system to achieve position, velocity, and direction [14]. This is a common

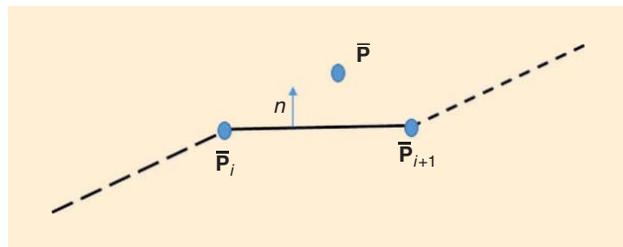


FIGURE 5. The road segment i with normal vector \hat{n} and a position estimate \bar{p} . The scalar product between the vectors can be used to determine whether or not the estimate should be considered to belong to the segment.

approach for military aircraft navigation and underwater navigation. It is possible to use this dead reckoning together with map matching to mitigate sensor imperfections. This is probably best achieved using dynamic filtering and will be described more in the sequel. For cheap commercial sensors, usually a GPS sensor is needed to handle the drift due to small sensor biases.

Dead reckoning: Map matching

As discussed previously, map matching can be done by fitting an estimate to the closest road or by looking at segments, etc. Here we will focus on the point-to-point matching, i.e., that the estimate is mapped to the closest orthogonal distance. In Figure 5, a position estimate (\bar{p}) is considered to belong to road segment i , i.e., between the two road edges \bar{p}_i and \bar{p}_{i+1} . This can easily be verified if the following scalar products are greater than zero: $\bar{p}_i \cdot \Delta_i > 0$ and $\bar{p}_{i+1} \cdot (-\Delta_i) > 0$, where $\Delta_i = \bar{p}_{i+1} - \bar{p}_i = (dX, dY)^T$. If this is the case, the closest distance to the segment can be calculated as $d = \|(\bar{p} - \bar{p}_i) \cdot \hat{n}\|$, using the normal vector $\hat{n} = (dY, -dX)^T$. Such a project is needed in many localization systems and, in particular, utilized in the method described in the section “Bayesian Filtering for Map-Aided Positioning.”

Dead reckoning: Dynamic filtering

A generic nonlinear filter for localization consists of the following main steps:

- *Time update:* Use a motion model to predict where the vehicle will be when the next measurement arrives.
- *Measurement update:* Use the current measurement and a sensor model to update the information about the current location.

In a Bayesian framework, the information is represented by the posterior distribution given all available measurements. The process of computing the Bayesian posterior distribution is called *filtering*. Details are given in the section “Bayesian Filtering for Map-Aided Positioning,” where the distance calculated in the section “Dead Reckoning: Map Matching” is used in a probabilistic way.

Bayesian filtering for map-aided positioning

Bayesian filtering

Nonlinear filtering is the branch of statistical signal processing concerned with recursively estimating the state x_t based on the measurements up to time t , $\mathbb{Y}_t \triangleq \{y_1, \dots, y_t\}$ from sample 1 to t . The most general problem it solves is to compute the Bayesian conditional posterior density $p(x_t | \mathbb{Y}_t)$. There are several algorithms for computing the posterior density. The KF [5] solves the filtering problem in case the model is linear and Gaussian. The solution involves propagating the mean $\hat{x}_{t|t}$ and the covariance $P_{t|t}$ for the posterior distribution. For nonlinear problems, the model can be linearized before the KF technique is applied, leading to the extended KF (EKF) [6]. Methods also exist where the Gaussian approximation is the key element, hence no linearization is needed; for instance, the unscented

Algorithm 1. The PF for map-aided positioning.

Given the system

$$x_{t+1} = f(x_t) + w_t$$

$$y_t = h(x_t) + e_t$$

- 1: Initialization: For $i = 1, \dots, N$, $x_{0|i-1} \sim p_{x_0^0}(x_0)$ and set $t = 0$.
- 2: PF measurement update: For $i = 1, \dots, N$, evaluate the importance weights $\tilde{\omega}_t^{(i)} = p(y_t | x_{t|i}^{(i)}, \mathbb{Y}_{t-1})$, and normalize $\omega_t^{(i)} = \tilde{\omega}_t^{(i)} / \sum_j \tilde{\omega}_t^{(j)}$ using map information.
- 3: Resample N particles with replacement:
 $\Pr\{x_{t+1}^{(i)} = x_{t|i}^{(i)}\} = \omega_t^{(i)}$.
- 4: PF time update: For $i = 1, \dots, N$ predict new particles
 $x_{t+1|i}^{(i)} \sim p(x_{t+1|i} | \mathbb{X}_t^{(i)}, \mathbb{Y}_t)$.
- 5: Increase time and repeat from step 2.

KF [36] approximates the posterior at each step with a Gaussian density. Common for these methods is that it is not trivial to impose hard constraints from the road map. They also do not work particularly well unless the posterior density is very monomodal or Gaussian. For a KF-based estimation with map information, see [37].

The road constraints imply a kind of information that normally leads to a multimodal posterior density (the target can be on either this road, or another road on the given map, etc.). Hence, a Gaussian approximation of the probability density function (PDF) is not suitable. A completely different approach to nonlinear filtering is based on approximating the posterior $p(x_t | \mathbb{Y}_t)$ numerically. The point mass filter (PMF) [38] represents the state space using a regular grid of size N , where the grid points and the related weights $(x^{(i)}, w_t^{(i)})$ are used as a representation of the posterior. Different basis functions have been suggested, the simplest one being an impulse at each grid, when the posterior approximation can be written $p(x_t | \mathbb{Y}_t) \approx \sum_{i=1}^N w_t^{(i)} \delta(x_t - x_t^{(i)})$, where $\delta(x)$ denotes the Dirac-delta function. The PF [10] is the state-of-the-art numerical solution today. It uses a stochastic grid $\{w_t^{(i)}, x_t^{(i)}\}_{i=1}^N$ that automatically changes at each iteration.

Depending on the model, it is also possible to implement numerical efficient filters combining KF and PF. The idea is to divide the state space into two parts. If there is a conditionally linear Gaussian substructure with this partition, the KF can be utilized for that part and the PF for the other part. This is referred to as the RBPF or the MPF [7]–[9], [39]–[41]. The RBPF improves the performance when a linear Gaussian substructure is present, e.g., in various map-based positioning applications and target tracking applications as shown in [41]. The map-aided positioning algorithm based on the PF is summarized in Algorithm 1.

PF-based map-aided positioning

In this section, the map-aided positioning method is first illustrated on experimental data. Then the crucial map-based observation is described in detail. Finally, the algorithm per-

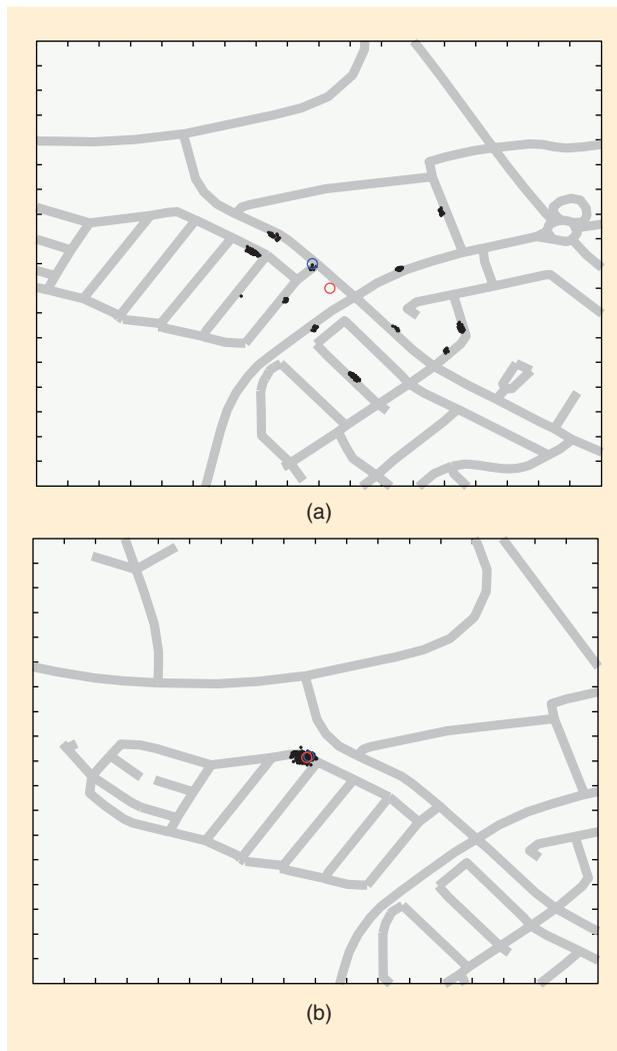


FIGURE 6. Map-aided positioning using WSS information in combination with road map information. The small black dots are particles, the red and blue circles represent mean estimate and ground truth (GPS) position, respectively. Represented are time instances after initialization when the filter is still in (a) a multimodal state and (b) at convergence. (a) The illustration of the particle cloud after some iterations. The multimodal PDF representing the position (several clusters of particles). The particles clustered but still the mean point estimate (red circle) does not correspond to the ground truth (blue circle). (b) The PF has converged to a unimodal PDF (one cluster). The mean estimate is now close to the ground truth.

formance is presented on ten experiments conducted in the same driving scenario.

Map-aided positioning illustrations

Figure 6 demonstrates the map-aided positioning using wheel speed information and road map information, where GPS information is used as a ground truth reference only; for other map-aided positioning applications, see, e.g., [42]–[48]. First, the PF is initialized in the vicinity of the GPS position. The initial distribution is chosen uniformly on road segments in a region around the GPS fix. Particles are allowed slightly off-road to handle off-road situations and small map errors. In Figure 6(a), the algorithm has been active for some time. As can be



FIGURE 7. Ten routes were driven (red GPS line), and the position using map-aided position is depicted for one of them (blue line).

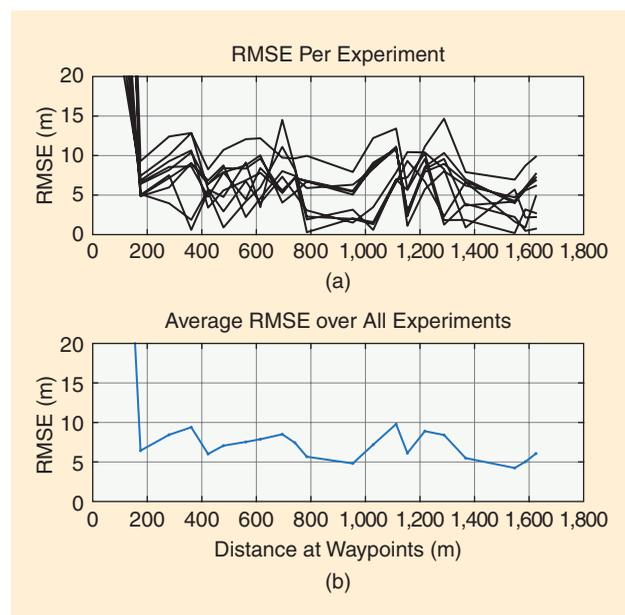


FIGURE 8. The RMSE for each trajectory as a function of driven distance evaluated at (a) specified waypoints and (b) as a total RMSE average at the same instances. Note that GPS was not used in the estimations but only as a source for ground truth evaluation.

seen, the PDF is highly multimodal (several clusters of particles). Note that the PF algorithm uses only wheel speeds from the CAN bus and that the GPS is only used to evaluate the ground truth. After some turns, the filter has converged and the mean estimate (red circle) is close to the true position (blue circle); see Figure 6(b).

Map-aided positioning algorithm

As discussed in the section “Dead Reckoning: Map Matching,” map matching can be used to fit an estimate to the closest road segment. In this section, we will focus on the PF implementation, so for each particle it is crucial to find the

closest road segment. The generic PF algorithm is used for map-aided positioning. The road map is used as a virtual sensor, so there is not an actual measurement function. Instead, the closest distance to every road segment is evaluated for each particle. The main advantage here compared to normal map-matching algorithms is that the entire probability density is considered, not just one point estimate. In Figure 5, a particle (\bar{p}) is considered to belong to road segment i , if the two defined scalar products are positive (see the section “Dead Reckoning: Map Matching”). The calculated distance can then be assumed Gaussian distributed and used in the PF measurement update.

To make the algorithm efficient, each particle will remember the road segment that was closest in the last update. Hence, if the distance is still close enough, not every road segment needs to be considered.

Map-aided positioning performance

To evaluate the average performance, ten similar experiments were conducted (see Figure 7). All trajectories are driven approximately in the same way, however, it is not possible to have them synchronized in time. Hence, root mean square error (RMSE) evaluation is done at some fixed waypoints. In Figure 8(a), the RMSE is calculated for each trajectory against known GPS-waypoints. In Figure 8(b), the average RMSE is presented; the performance is comparable to standard GPS position error.

Multilevel positioning and indoor navigation

Indoor navigation is challenging since no GPS signal is available. However, for parking garages, the map-aided positioning principle can still be utilized if maps are available. These maps can also include obstacles such as pillars, side walls, etc. defining boundary regions. Since the geometry and curvature are quite different than when driving on normal roads, it is also beneficial to extend the simple vehicle model and include a more precise spatial vehicle model. This can easily be done by evaluating side walls and pillars surrounding a rectangle placed around the particles that correspond to the vehicle’s geometry.

To handle multilevels, one successful implementation is to utilize entry and exit points between levels and simply handle each level as a continuation of the previous. Inertial sensors such as inclination in longitudinal accelerometer signal can also support the level-change decision. Similar geometries can be found in multilevel highways in large cities.

For the positioning to work well on highways, there must be some variation in the geometry. For instance, driving on very straight highways for a long time will, of course, lead to a more uncertain position estimate, since the wheel radius cannot be known or estimated exactly. However, for such scenarios it is very likely that GPS reception is very good and could be used to adjust the position estimate.

Measurement features

The PF method for positioning is very general, and it is easy to add other information sources to the measurement update.

There are several important and common information sources for automotive positioning that can be utilized. Modern camera-based vehicles can be equipped with traffic sign recognition, etc. If a database of the sign location is available, it fits perfectly into the map-aided positioning framework by simply adding a measurement update. Other information sources, such as maximum allowed speed information for a road, can be used for making a better proposal in the PF. In the future, more localization sources will probably be available, e.g., information exchange from the infrastructure.

Inertial sensors and velocity estimation

In this section, the use of inertial sensors such as accelerometers and gyros, both for dead reckoning, sensor fusion, and as stand-alone velocity estimation is discussed.

Inertial sensors

An IMU measures acceleration and angular rotation. From these measurements it is possible to integrate the underlying system to achieve an estimate of position, velocity, and direction [14].

Dead-reckoning IMU

The pure integration or dead-reckoning approach relies on very accurate and expensive sensors that are not possible to utilize in commercial passenger vehicles. The main problem for affordable commercial sensors is that they have both unknown sensor biases and drifts that are impossible to remove. For instance, in the longitudinal acceleration direction, both sensor errors and hills will act like an unknown time-varying bias. Integrating twice to achieve position will very rapidly yield large position errors. For reliable stand-alone navigation, without relying on, for instance, GPS, this is very difficult.

Map-aided positioning using IMU

An IMU-based map-aided position (stand-alone or in combination with WSS) was tested in [46], where the outcome was that the WSS is a superior velocity sensor due to the aforementioned problems.

GPS and IMU fusion

For modern ECUs, the in-vehicle sensor cluster consisting of rate gyros and accelerometers can be used together with a GPS sensor to achieve position, velocity, and orientation estimates [14]. However, not all vehicles have these signals, or they are not readily available. Hence, for many vehicles, application external sensors might be helpful, particularly utilizing signals available in smart phones.

Consider the following state vector

$$x = (q \ p^i \ v^b \ a^b)^T,$$

where q is the quaternion vector, p^i is the inertial position vector, v^b is the velocity in the body system, and a^b is the accelerometer vector in the body system. Let C_{ib} be the conversion matrix from body to inertial systems. Hence, in dis-

crete time, we have the inertial position $p_{i+1}^i = p_i^i + TC_{ib}v_i^b$. It can be shown that

$$\dot{q}_{bi} = \frac{1}{2}S(\omega_{bi}^b)q_{bi},$$

where ω_{bi}^b is the angular velocity of the body system relative to the inertial system described in body coordinates.

Figure 9 shows the EKF estimate based on IMU and GPS signals from a Google Nexus mobile phone using the data logging from [49], together with the GPS position using a discretized model. For this application, it is essential that some velocity estimation algorithm or position sensor is utilized to mitigate the dead-reckoning problem with unknown biases in accelerometer and gyro signals. Here the GPS sensor was utilized instead of map-matching techniques.

Virtual speedometer

For some applications, an accurate velocity estimate is essential. A complementary method to the previously described estimation is to utilize frequency analysis. It turns out that the velocity of the vehicle is proportional to vibrations in the accelerometer signal [50]. This can be utilized in the previously described positioning filter. It is not as accurate as WSS information, but for a stand-alone application when WSS is not available, data from a standard smart phone can be utilized. In Figure 10, the lateral accelerometer spectrogram is depicted together with the angular velocity from the WSS (as ground truth). The spectrum is formed at every instance (downsampled to every second) by filtering the periodogram of the lateral acceleration. There is usually a frequency related to the velocity. Note that there are usually some overtones as well. Utilizing this in the frequency domain, it is possible to construct a simple peak detector to estimate the velocity. The point estimates utilizing only the maximal peak in the accelerometer spectrum (batch-wise every second) is depicted

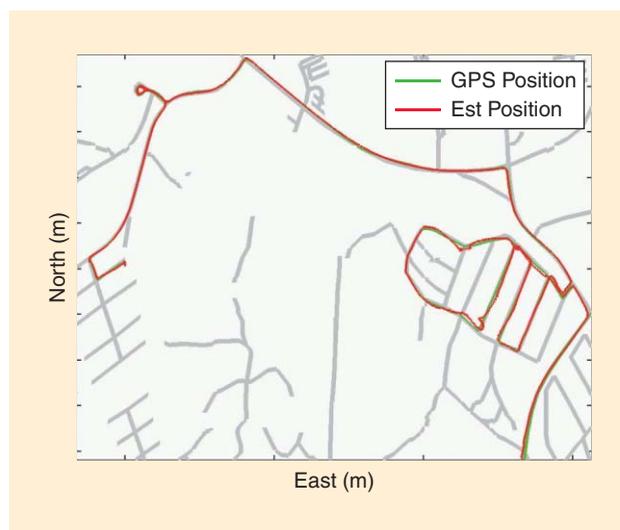


FIGURE 9. The fused position EKF estimate and the GPS position depicted in a street map utilizing sensor data from a Google Nexus Android mobile phone.

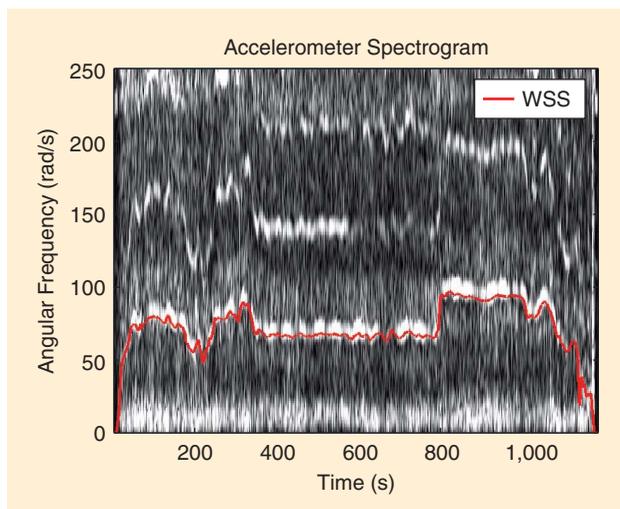


FIGURE 10. An accelerometer spectrogram (filtered periodogram of the lateral acceleration as a function of time) compared to the velocity from the WSS, where the harmonics correspond to the wheel rotation frequency.

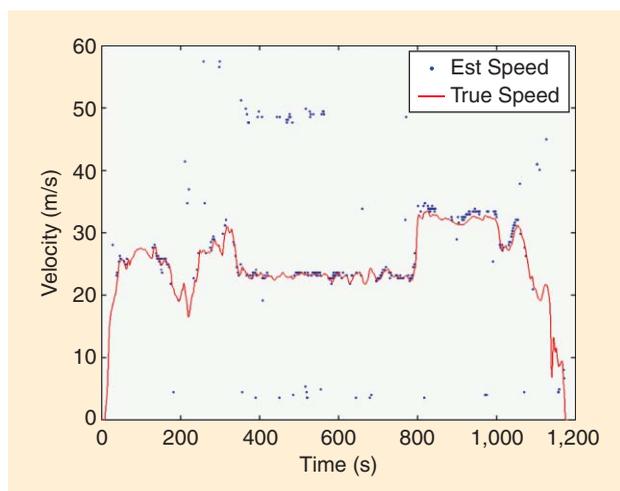


FIGURE 11. Velocity from the WSS compared to estimates derived from peaks in the accelerometer spectrum. Sometimes there are outliers or the wrong harmonic is selected, however, this can later be corrected by the overall velocity filter.

in Figure 11. Usually this correlates to the true velocity, but there are some outliers. It is possible to improve the detector by incorporating the knowledge of overtones, but this is not done in this article.

For map-aided positioning, WSS information is always available. But for stand-alone applications, the vibration-based velocity estimate could be used as a complementary measurements for map-aided positioning. Figure 12 illustrates how the speed estimation can be used as input signal (instead of the WSS data) to the map-aided positioning. For low velocities, this is not accurate enough, but pure accelerometer integration can potentially be used for short periods of time to support the algorithm.

Discussion and conclusions

We have discussed the needs in future automotive localization algorithms and pointed out that both accuracy and integrity

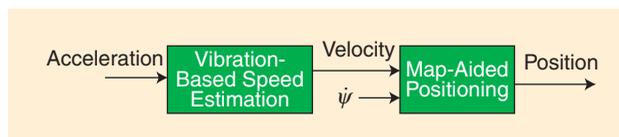


FIGURE 12. The general idea for stand-alone map-aided positioning when the WSS is not available is to use the acceleration vibration speed estimation as input to the map-aided positioning algorithm. With proper outlier rejection to the velocity estimates, the algorithm can function without WSS data.

have to be improved compared to the navigation systems today that rely on GNSS. For this purpose, we outlined a path to future automotive localization algorithms based on a statistical signal processing approach, where information from various sensors and information sources are fused based on given sensor models and an odometric motion model. The possible sensors include in-vehicle sensors such as WSS, accelerometers, gyros, and external ones such as GPS. However, localization concerns the relative position of the own vehicle compared to the surrounding, so the position relative the road network is more important than the absolute longitude and latitude. A road map is the key information source for this purpose, and we have discussed the concepts of map matching (basically projection of a position to the road network) and map-aided positioning (where the road map is treated as a sensor). Furthermore, landmarks such as road signs detected by a camera and the inclusion of car to infrastructure information, and wireless sources (Bluetooth, Wi-Fi, and mobile positioning) will be crucial in the future, and this information is also easily incorporated in our framework.

We have, in particular, highlighted the crucial concept of map-aided positioning. Utilizing measurements from a yaw rate and wheel speed signals, we have shown that it is possible in urban areas to position a vehicle with almost GPS accuracy without using any external GNSS positioning sensor when utilizing the vehicle model and accurate road map information in a PF. The incorporation of inertial sensor measurements for velocity estimation utilizing accelerometer vibrations was demonstrated, as well as the basic principle when using it for positioning.

Acknowledgment

This work was partially supported by the Wallenberg Autonomous Systems Program.

Authors

Rickard Karlsson (rickard@isy.liu.se) received his M.Sc. degree in applied physics and electrical engineering in 1996 and his Ph.D. degree in automatic control in 2005, both from Linköping University, Sweden, where he is currently an associate professor. He has worked with automotive signal processing applications at NIRA Dynamics since 2007 and target tracking applications at Saab Dynamics in Linköping between 1997 and 2002. His research interests include positioning and tracking applications mainly using particle filters.

Fredrik Gustafsson (fredrik@isy.liu.se) received his M.Sc. degree in electrical engineering in 1988 and his Ph.D. degree in

automatic control in 1992, both from Linköping University, Sweden. He is a professor of sensor informatics in the Department of Electrical Engineering, Linköping University. His research is focused on sensor fusion and statistical methods in signal processing, with applications to aerospace, automotive, audio, and communication systems. He is the author of four books, more than 100 international papers, and 14 patents. He is also a cofounder of three spin-off companies in these areas. He is an associate editor of *IEEE Transactions on Signal Processing*.

References

- [1] DARPA. (2016). The grand challenge and urban challenge. [Online]. Available: <http://archive.darpa.mil/grandchallenge/>
- [2] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, M. Sokolsky, G. Stanek, D. Stavens, A. Teichman, M. Werling, and S. Thrun, "Towards fully autonomous driving: Systems and algorithms," in *Proc. IEEE Intelligent Vehicles Symp. (IV)*, 2011, pp. 163–168.
- [3] Google. (2016). Google: The self-driving car. Available: <http://www.google.com/selfdrivingcar>
- [4] K. Jo, J. Kim, D. Kim, C. Jang, and M. Sunwoo, "Development of autonomous car part i: Distributed system architecture and development process," *IEEE Trans. Ind. Electron.*, vol. 61, no. 12, pp. 7131–7140, Dec. 2014.
- [5] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. Am. Soc. Mech. Eng. J. Basic Eng.*, vol. 82, pp. 35–45, Mar. 1960.
- [6] S. F. Schmidt, "Application of state-space methods to navigation problems," *Adv. Control Syst.*, vol. 3, pp. 293–340, 1966.
- [7] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. Comput.*, vol. 10, pp. 197–208, 2000.
- [8] A. Doucet, N. J. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 613–624, Mar. 2001.
- [9] R. Chen and J. S. Liu, "Mixture Kalman filters," *J. Royal Stat. Soc. Series B (Stat. Methodol.)*, vol. 62, no. 3, pp. 493–508, 2000.
- [10] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEEE Proc. F Radar Signal Process.*, vol. 140, no. 2, pp. 107–113, Apr. 1993.
- [11] I. Skog and P. Händel, "In-car positioning and navigation technologies: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 1454–1469, Mar. 2009.
- [12] W. J. Fleming, "Overview of automotive sensors," *IEEE Sensors J.*, vol. 1, no. 4, pp. 296–308, Dec. 2001.
- [13] W. J. Fleming, "New automotive sensors: A review," *IEEE Sensors J.*, vol. 8, no. 11, pp. 1900–1921, Nov. 2008.
- [14] D. Titterton and J. Weston, *Strapdown Inertial Navigation Technology*, 2nd ed. Stevenage, U.K.: IET, 2004.
- [15] "ESRI shapefile technical description: An ESRI white paper," July 1998. [Online]. Available: <https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>
- [16] Open street map [Online]. Available: <http://openstreetmap.org>
- [17] S. Ammoun, F. Nashashibi, and A. Brageton, "Design of a new GIS for ADAS oriented applications," in *Proc. IEEE Intelligent Vehicles Symp. (IV)*, June 2010, pp. 712–716.
- [18] M. Haklay and P. Weber, "Openstreetmap: User-generated street maps," *Pervasive Comput.*, vol. 7, no. 4, pp. 12–18, Oct. 2008.
- [19] C. Whitea, D. Bernstein, and A. Kornhausera, "Some map matching algorithms for personal navigation assistants," *Transport. Res. Part C: Emerg. Technol.*, vol. 8, no. 1–6, pp. 91–108, 2000.
- [20] J. Du and M. Barth, "Bayesian probabilistic vehicle lane matching for link-level in-vehicle navigation," in *Proc. IEEE Intelligent Vehicles Symp.*, June 2006, pp. 522–527.
- [21] D. Obradovic, H. Lenz, and M. Schupfner, "Fusion of sensor data in Siemens car navigation system," *IEEE Trans. Veh. Technol.*, vol. 56, no. 1, pp. 43–50, 2007.
- [22] M. A. Qudus, W. Y. Ochieng, and R. B. Noland, "Current map-matching algorithms for transport applications: State-of-the art and future research directions," *Transport. Res. Part C: Emerg. Technol.*, vol. 15, no. 5, pp. 312–328, 2007.
- [23] O. Noren, "Monitoring of road surface conditions," M.S. thesis, Dept. Electrical Engineering, Linköping Univ., Sweden, 2014.
- [24] Y. Cui and S. S. Ge, "Autonomous vehicle positioning with GPS in urban canyon environments," *IEEE Trans. Robot. Automat.*, vol. 19, no. 1, pp. 15–25, Feb. 2003.
- [25] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping (SLAM): Part I," *IEEE Robot. Automat. Mag.*, vol. 13, no. 2, pp. 99–110, June 2006.
- [26] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): Part II," *IEEE Robot. Automat. Mag.*, vol. 13, no. 3, pp. 108–117, Sept. 2006.
- [27] S. Thrun, W. Burgard, and D. Fox, "Probabilistic robotics," in *Ser. Intelligent Robotics and Autonomous Agents*. Cambridge, MA: MIT Press, 2005.
- [28] M. Montemerlo, S. Thrun, D. Koller, and B. Wegbreit, "FastSLAM: a factored solution to the simultaneous localization and mapping problem," in *Proc. AAAI National Conf. Artificial Intelligence*, Edmonton, Canada, 2002.
- [29] J. Levinson and S. Thrun, "Map-based precision vehicle localization in urban environments," in *Proc. Robotics: Science and Systems*, 2007.
- [30] N. Mattern, R. Schubert, and G. Wanielik, "High-accurate vehicle localization using digital maps and coherency images," in *Proc. IEEE Intelligent Vehicles Symp. (IV)*, 2010, pp. 462–469.
- [31] Y. Zhao, *Vehicle Location and Navigation Systems*. London: Artech House, 1997.
- [32] F. Gustafsson, F. Gunnarson, N. Bergman, U. Forsell, J. Jansson, R. Karlsson, and P. J. Nordlund, "Particle filters for positioning, navigation, and tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 425–437, Feb. 2002.
- [33] F. Gustafsson, U. Orguner, T. B. Schön, P. Skoglar, and R. Karlsson, "Navigation and tracking of road-bound vehicles," in *Handbook of Intelligent Vehicles*. London: Springer-Verlag, 2012, pp. 397–434.
- [34] M. Ulmke and W. Koch, "Road-map assisted ground moving target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 4, pp. 1264–1274, Oct. 2006.
- [35] D. Salmond, M. Clark, R. Vinter, and S. Godsill, "Ground target modelling, tracking and prediction with road networks," in *Proc. Int. Conf. Information Fusion*, July 2007.
- [36] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proc. American Control Conf.*, Seattle, WA, June 1995, pp. 1628–1632.
- [37] K. Jerath and S. Brennan, "GPS-free terrain-based vehicle tracking on road networks," in *Proc. American Control Conf.*, June 2012, pp. 307–311.
- [38] S. C. Kramer and H. W. Sorenson, "Recursive Bayesian estimation using piece-wise constant approximations," *Automatica*, vol. 24, no. 6, pp. 789–801, Nov. 1988.
- [39] C. Andrieu and A. Doucet, "Particle filter for partially observed Gaussian state space models," *J. Royal Stat. Soc. Series B (Stat. Methodol.)*, vol. 64, no. 4, pp. 827–836, 2002.
- [40] T. Schön, F. Gustafsson, and P. J. Nordlund, "Marginalized particle filters for mixed linear/nonlinear state-space models," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2279–2289, July 2005.
- [41] T. B. Schön, R. Karlsson, and F. Gustafsson, "The marginalized particle filter in practice," in *Proc. IEEE Aerospace Conf.*, Big Sky, MT, Mar. 2006.
- [42] P. Hall, "A Bayesian approach to map-aided vehicle positioning," M.S. thesis, Dept. Electrical Eng., Linköping Univ., Sweden, 2001.
- [43] Y. Cheng and T. Singh, "Efficient particle filtering for road-constrained target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 4, pp. 1454–1469, Oct. 2007.
- [44] N. Svenzén, "Real time map-aided positioning using a Bayesian approach," M.S. thesis, Dept. Electrical Eng., Linköping Univ., Sweden, 2003.
- [45] J. Kronander, "Robust vehicle positioning: Integration of GPS and motion sensors," M.S. thesis, Dept. Electrical Eng., Linköping Univ., Sweden, 2004.
- [46] G. Hedlund, "Map aided positioning using an inertial measurement unit," M.S. thesis, Dept. Electrical Engineering, Linköping Univ., Sweden, 2008.
- [47] P. Davidson, J. Collin, J. Raquet, and J. Takala, "Application of particle filters for vehicle positioning using road maps," in *Proc. 23rd Int. Tech. Meeting Satellite Division Institute Navigation*, Sept. 2010, pp. 1653–1661.
- [48] P. Davidson, J. Collin, and J. Takala, "Application of particle filters to a map-matching algorithm," *Gyrosc. Navigat.*, vol. 2, no. 4, pp. 285–292, 2011.
- [49] G. Hendeby and F. Gustafsson. (2013). Sensor fusion app: Sensor fusion at Linköping Univ. [Online]. Available: <https://play.google.com/store/apps/details?id=com.hiq.sensor>
- [50] R. Karlsson and F. Gustafsson, "Velocity estimation (n2515-028-wop00)," Patent application filed, Aug. 2015.

DSP Applications in Engine Control and Onboard Diagnostics

Enabling greener automobiles

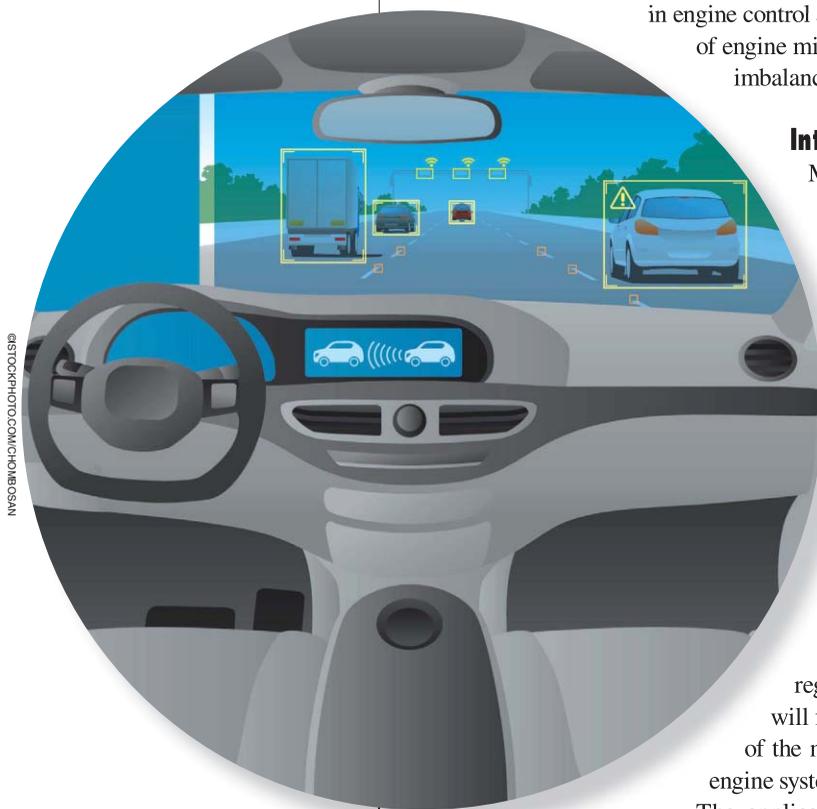
Digital signal processing (DSP) has been playing an increasingly important role in engine control and onboard diagnostics (OBD), a critical area of vehicle powertrain controls, to meet increasingly strict fuel efficiency requirements and emission regulations. This article gives an overview of DSP applications in this field from a practical perspective and identifies fruitful areas for DSP research and development in engine control and OBD. It also provides examples in the areas of engine misfire detection, individual cylinder fuel-air ratio imbalance detection, and engine knock detection.

Introduction

Microprocessor-based electronic systems have become prevalent in modern automobiles, as they make it possible to meet increasing demands in powertrain control, chassis control, OBD, driver assistance, and numerous other emerging applications. DSP represents a predominant form of computation being performed by these microprocessors and has an important role in many automobile applications.

DSP has become indispensable for many advanced engine control and diagnostic functions because signal processing makes it possible to implement the control and diagnostics needed to meet increasingly strict fuel efficiency requirements and emission regulations. Throughout this article, the discussion will focus on gasoline engine applications, but many of the methods outlined also apply to diesel and other engine systems.

The applications of signal processing to engine control essentially started with the advent of microprocessor-based engine control units (ECUs) in the 1980s. Over time, the complexity of the associated DSP algorithms has evolved from simple to sophisticated processing, much in line with the increase in ECU processing capabilities. The application of signal processing started from the exponential moving average filter, which is used for signal conditioning and smoothing because of its simplicity. The microprocessors or microcontrollers used in modern ECUs have evolved, and have significantly more computational power,



©ISTOCKPHOTO.COM/CHIRIBOSAN

with some high-end microprocessors having dedicated built-in DSP functions, which significantly speed up signal processing and increase the computational efficiency. The result is that not only are finite impulse response (FIR) and infinite impulse response (IIR) filters quite feasible on modern platforms, but considerably more complex algorithms, such as fast Fourier transform (FFT), discrete Fourier transform (DFT), neural networks, and even modern machine-learning algorithms, have become possible.

In the automotive industry, the development of engine control and diagnostics is outpacing the development of the engine “hardware” itself. Sophisticated algorithms based on modern control theory and modern signal processing are being integrated into engine control and OBD because simple and intuitive algorithms no longer satisfy stringent fuel economy and emission requirements. The deployment of these advanced algorithms is facilitated by many new sensors being added to engine systems. The end result is that an abundance of raw signal data is available for processing, and DSP is essential for processing and cleansing this data, performing signal detection and important parameter estimation needed for vehicular processes. In particular, the automobile is a rotational machine and, as such, produces signals with periodic characteristics, and DSP techniques, such as frequency domain filtering, are both effective and essential to

processing the data so that engine control and diagnostics challenges can be addressed in a robust manner.

This article provides a survey of DSP applications in engine control and OBD with a focus on practical automotive applications. The application examples we explore will mainly include:

- engine misfire detection and related resonance noise removal and rough road detection
- individual cylinder fuel-air ratio imbalance detection for control and diagnostics
- engine knock detection.

Our discussion will cover many DSP methods used in the automotive industry, ranging from traditional to advanced signal processing techniques. The overview will focus primarily on the most common methods currently being applied to solve real problems faced by automotive manufactures. The discussion will conclude with a summary of DSP applications in the automotive industry and highlight practical implementation matters. Figure 1 provides an overall taxonomy of key application categories and techniques reviewed in this article.

OBD

OBD has become increasingly essential in modern automobiles. Today, all light-duty vehicles are required to be equipped with OBD systems by the California Air Resources

Analysis Domain	Crankshaft Angular Position	Oxygen Sensor	Knock Sensor	Cylinder Pressure Sensing	Ion Sensing
Time-Domain		X	X		
Crank Angle Domain	X	X			
Frequency Domain	X	X	X		
Techniques					
Filtering-Based Methods	X	X	X		
Observers	X	X			
Statistical Techniques	X				X
Model-Based Methods	X	X			
Neural Networks	X				X

FIGURE 1. The taxonomy of DSP applications in engine controls and diagnostics.

Board (CARB). Various OBD systems monitor the performance of virtually every component that can affect the emission performance of the vehicle. If a fault is detected by an OBD system, a warning indicator or lamp on the vehicle instrument panel will illuminate to alert the driver.

A general OBD system is illustrated in Figure 2. The OBD system starts by collecting the signals to be analyzed, a detection and decision-making system, a bookkeeping system, and a malfunction indicator light. The detection and decision-making system is the core of the OBD system and carries out the detection of signals or features of interest and is responsible for making a decision as to whether a monitored device or system has failed. The bookkeeping system is responsible for interpreting the output of the detection system and deciding whether and how to illuminate the warning light based on regulation rules. Among these subsystems, the detection system is the most challenging for manufacturers to design because the detection accuracy must be very high to meet CARB's requirements, while the false alarm rates must also be very low to minimize the warranty cost caused by any false alarms.

Engine misfire detection

The misfiring of an automobile engine can directly lead to increased emissions into the atmosphere and cause potential damage to the catalyst. The engine misfire monitor is one of the most challenging OBD systems to design. Numerous misfire detection methods have been researched, published, and patented by automotive companies, suppliers, universities, and research institutes. Even today, achieving the full range and robust misfire detection performance needed, while keeping implementation costs low, is a very challenging task faced by automotive manufacturers, especially as there are a variety of new, emerging engine configurations and control strategies being deployed to enhance fuel economy.

The main challenge in the development of an on-board real-time misfire detector is finding a unified and robust algorithm that can detect misfires and faulty cylinders with high detection accuracy, while having little to no false alarms, under all of the required engine operating conditions, misfire patterns, and potential sources of noise interference. The engine operating conditions include various engine speeds and engine loads, engine cold start, gear positions, mechanical dynamics, dynamic skip firings, and many other conditions. The strictest requirement for engine operating conditions is maintaining a full range of misfire detection capabilities that cover the

complete range of engine speeds and load conditions, with the notable exception at high engine speed and low load corner that are defined by CARB. The underlying misfire patterns can be quite varied, ranging from regular to random to having special sequence characteristics. Further, misfires may also occur in paired or multiple cylinders. As an example of the challenge faced by misfire detectors, a recently proposed dynamic skip firing control strategy selectively deactivates cylinders to match the torque demand to improve fuel economy [1], and this new algorithm poses new challenges to some misfire detection methods because of the potential to inaccurately distinguish misfires from skip firing. Further challenging misfire detection is noise interference that may arise from various sources, such as, powertrain resonance, rough road surfaces, and electric motor torque mixed with engine torque in electrical hybrid vehicles.

Because of the difficulties in meeting regulation requirements, automotive engineers have intensely studied a wide variety of engine signals, with engine speed or crankshaft speed fluctuation becoming the mainstream since onboard misfire monitors were required on automobiles, as illustrated by two decades of literature, such as those in [1]–[12]. The main reasons that methods based on engine speed have become mainstream in the automotive industry are the easy availability of the underlying signal, overall low implementation costs, and adequate detection performance under most conditions. With a good design that uses signal processing techniques, it is possible to design a robust engine-speed-based misfire detection algorithm that can meet CARB's current full range of detection requirement.

In recent years, ion signals have been increasingly studied for misfire detection [13], [14], but misfire detection algorithms based on the ion signal currently are still only rarely used in mass production because of the immaturity of the technology, the current costs, and other constraints. Bringing the cost down and ensuring reliable detection using ion signals is a promising direction for future signal processing research.

Beyond engine crankshaft speed and ion signals, other signals that have been explored in the past for supporting misfire detection include the signals from measurement of instantaneous net engine torque, combustion chamber pressure, exhaust gas pressure, and optical observation of gas flows, as well as structure-borne sound.

Misfire detection using crankshaft speed fluctuation has mainly involved signal processing algorithms and other intuitive algorithms that, for example, apply model-based methods or attempt engine roughness analysis. Model-based misfire detection methods stem from the physical relationship between the engine crankshaft speed and engine torque and cylinder pressures. When each cylinder in an internal combustion engine fires, a torque pulse is applied to the crankshaft throw. When a cylinder misfires, the torque applied to the crankshaft is altered, resulting in an altered engine and driveline angular velocity response [15]. The pressure profiles associated with misfire cylinders are quite different from those of normal firing cylinders, and thus misfire detection is easy if the engine

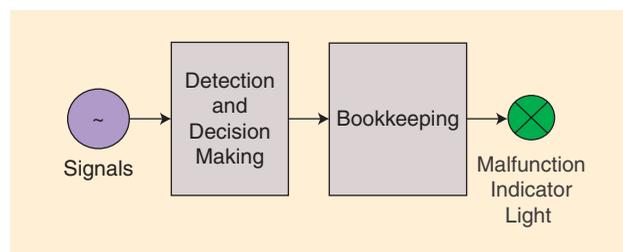


FIGURE 2. A block diagram of a general OBD system.

torque or cylinder pressures are available. Unfortunately, currently the installation of a torque sensor or individual cylinder pressure sensors in production vehicles for measuring engine torque or cylinder pressures is not practical due to their high cost and other technical concerns. Thus, many researchers have tried to establish various physical models to reconstruct torque or pressure profiles from the measured crankshaft speeds and, from these, make misfire decisions. An overview of the methods can be found in [15].

The problem with model-based methods is that it is very difficult to establish a robust model to meet the required misfire detection performance under all required engine operating conditions. Even with a very accurate model that has been well calibrated, the model-generated responses from the measured crankshaft speeds, which are commonly contaminated by various sources of noise and interference, may be severely distorted and not suitable for making accurate misfire decisions. In addition, model calibration is not a simple task for deployment in production vehicles, and thus complicated physical models were seldom used by automotive manufacturers for misfire detection.

To overcome the aforementioned problems, some researchers studied simplified models or solutions based on physical concepts. For example, the misfire detection is based on the indicated mean effective pressure (IMEP) estimated from crankshaft angular speeds [16], [17]. IMEP represents an average pressure of a cylinder during a combustion cycle, and it will decrease when a misfire occurs. Thus, misfire detection could be performed through the comparison of estimated IMEP of a cylinder with a preset threshold. In [16], a simple IMEP estimation model is derived that is proportional to the difference of the square of crankshaft angular speeds at the sample locations. A simpler approximation from experiments uses an estimate of IMEP proportional to the difference of crankshaft angular speeds [17].

Methods based on engine roughness analysis methods use the concept that the angular acceleration is proportional to the derivative of the engine rotational energy. Engine misfiring results in a decrease in engine crank rotational energy. A metric for engine roughness was introduced and able to be calculated in real time [2]. A misfire decision is made by comparing the calculated engine roughness with given thresholds in terms of engine speed and load. Because of its simplicity and relative low computational load, it was easily implemented in ECUs for complying with early OBD requirements [2]. However, it is nearly impossible to reliably work beyond an engine speed of 3,000 rev/min, not to mention the full range of detection requirements that were later imposed. To improve the detection performance, some developers have used low-pass filters and median

filters to reduce the noise existing in the angular acceleration calculated from engine speed or angular time intervals [3].

Firings of an engine are arranged periodically in every engine cycle for all engine cylinders. Many signatures embedded in crankshaft angular speed signals present deterministic periodic characteristics when they are viewed from the crankshaft angular domain. For example, the engine firing frequency is constant when the crankshaft angular speeds are sampled in the crankshaft angular domain. Thus, frequency-domain signal processing techniques can play a very useful and important role in the design of misfire detection algorithms.

A DSP approach to a misfire detection algorithm was presented in [4], in which the misfire detection system achieves a high degree of accuracy through the combination of multi-stage signal conditioning, multirate signal processing, statistical decision technology, and a mixed window size sampling strategy. This approach is capable of being installed on a large variety of motor vehicles through the use of existing on-board automotive microcontroller technology.

As illustrated in the block diagram for the misfire detection system in Figure 3 [4], which heavily uses signal processing, the engine crankshaft angular speed (labeled “RPM” in the figure) is its major input signal for misfire signatures in the system. Another input is the manifold absolute pressure (labeled “MAP” in the figure) indicating engine load. It uses a medium data rate (e.g., three data points per firing event for a four-cylinder engine) to avoid signal aliasing that could possibly occur in a low data rate (one data point per firing event) system and avoid the high computational requirements associated with a high data rate system. The signal preprocessing extracts average and fluctuation signal components from the engine speed and manifold absolute pressure signals using digital filters. The preprocessed engine speed and manifold absolute pressure signal components are used to compensate the influence of engine operating conditions through the process of misfire magnitude equalization. The equalized signal is then filtered to further enhance the signal-to-noise ratio (SNR) through the equivalent bandpass filtering. To reduce the computational load, the filtered signal is decimated into a

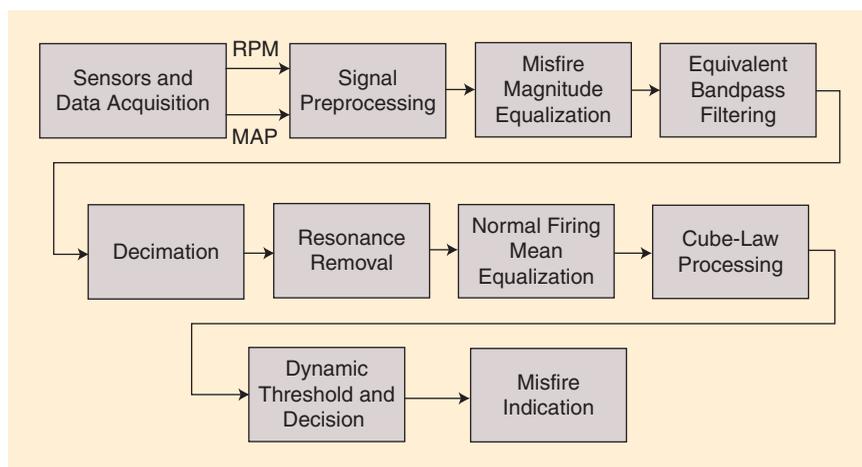


FIGURE 3. A DSP-based misfire detection system. (Figure adapted from [4].)

low data rate signal before removing resonance noise that is potentially caused by powertrain dynamics. The signal is then further processed through normal firing mean equalization, which results in a firing data point output signal with a zero mean. This signal is then subsequently processed with a cube law processing to further enhance the separation between misfire and normal firing signatures, and thus improve detection performance.

Unlike many conventional methods, this system does not use any look-up threshold table for misfire decisions. To achieve a robust misfire decision, instead, it makes misfire decisions through a dynamic threshold mechanism based on the signal's statistics information from previous normal firing and misfire signatures. The statistical information is obtained with the running mean and deviation estimators, which use low-pass filtering.

As shown in this example, signal processing directly leads to high performance in misfire detection, and signal processing further reduces system calibration efforts for various vehicle families because many filter parameters can be predetermined based on engine information, such as the number of cylinders.

Over the decades, many other signal processing methods have been explored for misfire detection, including FFT, time-frequency analysis, and wavelet methods [6]–[8]. The main challenges for FFT and spectrum methods are identifying the misfired cylinder, high data sampling rates, and high computational requirements. Wavelet methods significantly increase implementation complexity, with much higher computational requirements, and thus are not readily applicable for onboard misfire detection in today's production engine controllers. In addition, adaptive signal processing methods and Kalman filtering methods, as well as neural network techniques have been studied for misfire detection [9]–[12]. But while numerous methods have shown that they can detect misfires in some cases, only a few misfire detection algorithms have proven to robustly work across the full range of engine operating conditions required by regulations. Further, for an automotive manufacturer, the implementation cost is another critical consideration, should a method even meet the technical requirements. Ultimately, it is desirable to have a single algorithm that works across a broad spectrum of engines with different numbers of cylinders (e.g., ranging from two to ten cylinders) as this will reduce algorithm implementation cost and shorten the learning curve for calibration by engineers.

Resonance noise removal

In the misfire detection system discussed previously, there is a special processing to remove resonance noise. This kind of noise is difficult to effectively filter out or suppress using conventional filtering methods because it is a nonadditive noise with time-variant resonance frequencies and damping characteristics inherent in the crankshaft speed signals. It occurs usually at high (fourth or higher) transmission gear and low engine speed and is basically caused by the resonance characteristics of the powertrain dynamics when an impulse-type source excites the system. A large loss in engine

torque due to misfire excites the powertrain and generates damped oscillations in engine speed. If the resonance interference is strong enough, false misfire counts may become triggered, and the resulting increase in “false alarms” is a major concern underlying the cost-effectiveness of warranties for automobile manufacturers.

Based on this explanation for resonance noise, a signal processing technique using deconvolution is presented in [18] to effectively detect and remove the noise due to resonance by utilizing resonant and inverse filter banks. The misfire signal involved with resonance is modeled as the result of an impulse signal convolved with a resonant system, which can be described as a second order system. If the resonant frequency and damping factor are determined, then the misfire signal can be recovered from the deconvolution of the misfire signal corrupted by resonance with the resonant system, i.e., the convolution of the misfire signal with the inverse dynamics of the resonant system.

The resonance frequencies vary with engine operating conditions, and thus a resonant filter bank with different resonant frequencies is used to determine resonance frequencies. Each resonant filter in the bank has the same input and is followed by a square-law detector. The maximum output in amplitude from the square-law detectors is used to determine whether resonance noise exists and its associated frequency. At the same time, the same input signal for the resonant filters is input into each inverse filter. The output signal from the inverse filter corresponding to the resonant filter with the maximum output is the best resonance-removed signal and is taken as the output of the processing system. Figure 4 shows the resonance noise detection and removal system [18]. Another advanced approach for the resonant frequency determination and noise removal is to use an adaptive signal processing technique for achieving better performance, in which the resonance frequency is adaptively estimated and the coefficients of the inverse filter are updated with the estimated resonance frequency.

Rough road detection

In addition to challenges that low load and high engine speed, or engine speed data that are corrupted resonance interference, pose to misfire detection, interference associated with rough road conditions is another challenge for misfire detection algorithms. Signal processing can significantly suppress interference from rough road conditions in misfire detection and improve the misfire detection capability when a vehicle is driving on a rough road. For a robust DSP-based misfire detection algorithm, it is possible to continuously work without disabling misfire monitoring under rough road conditions. However, the current reality is that most misfire detection algorithms that use crankshaft speed fluctuation still suffer from rough road interference during misfire monitoring: either normal firings sometimes are falsely detected as “misfires” or real misfires sometimes fail to be detected. These algorithms have to disable misfire monitoring when the vehicle encounters rough road conditions, which is currently allowed by CARB regulations.

To disable misfire monitoring under rough road conditions, a rough road indication has to be determined in real time. The definition of roughness for rough roads is somewhat subjective, and the detection of rough road itself is not trivial. Popular rough road detection methods are based on signals from wheel speed sensors or crankshaft speed signals. Both left and right wheel speed signals are usually used as inputs in wheel speed methods [19]. The high-frequency components of the wheel signal contains the roughness signature of a road surface. First, the wheel speed signals are filtered with high-pass filters. The output of the high-pass filters are then squared and passed through low-pass filters for averaging. The square root of each averaged signal is then compared with a predetermined threshold to generate a roughness indicator flag for use in the misfire monitor.

An FFT-based rough road detector using engine crankshaft speed as input was explored in [20], and its system block diagram is illustrated in Figure 5. It first uses a filter to remove undesired components outside of a rough road frequency range. Then, the filtered signal is resampled for data reduction and rearranged for FFT. A power spectrum signal is calculated from the FFT components. The energy of the power spectrum signal is obtained by integrating the power of all its individual components over a narrow frequency band, which usually ranges from 1 to 5 Hz for rough road detection, depending on engine operating conditions, such as engine speed and load. A rough road is detected when the energy signal is greater than the energy threshold, which is a function of engine speed and load. A more advanced rough road detection method is presented in [21], which is based on adaptive signal processing and statistical analysis techniques and does not use wheel speed sensors, but instead uses engine crankshaft speed and transmission output speed as its inputs.

Individual cylinder fuel-air ratio imbalance detection for control and diagnostics

Engines produce drive torque to propel a vehicle through combustion of a mixture of fuel and air in cylinder chambers. A fuel-air ratio imbalance across cylinders could result in poor fuel economy and higher exhaust emissions. Thus, individual cylinder fuel-air ratio balance detection and control is an important feature in modern engine controls. CARB has also required automotive manufacturers to install fuel-air ratio imbalance on-board diagnosis systems in vehicles starting with the 2011 model year. Thus, both individual cylinder fuel-air ratio control and individual cylinder fuel-air ratio imbalance detection and monitoring are important for the improvement of fuel economy, reduction of exhaust emissions, and satisfaction of CARB's regulations.

The most critical element for proper control and diagnostics is the real-time retrieval of fuel-air ratio information associated with individual cylinders. Currently, regulations only require that each engine bank's fuel-air ratio imbalance is reported, instead of having each individual cylinder report its fuel-air ratio imbalance. However, information that can identify which cylinder is experiencing fuel-air ratio imbalance is extremely important for individual cylinder fuel-air ratio control because

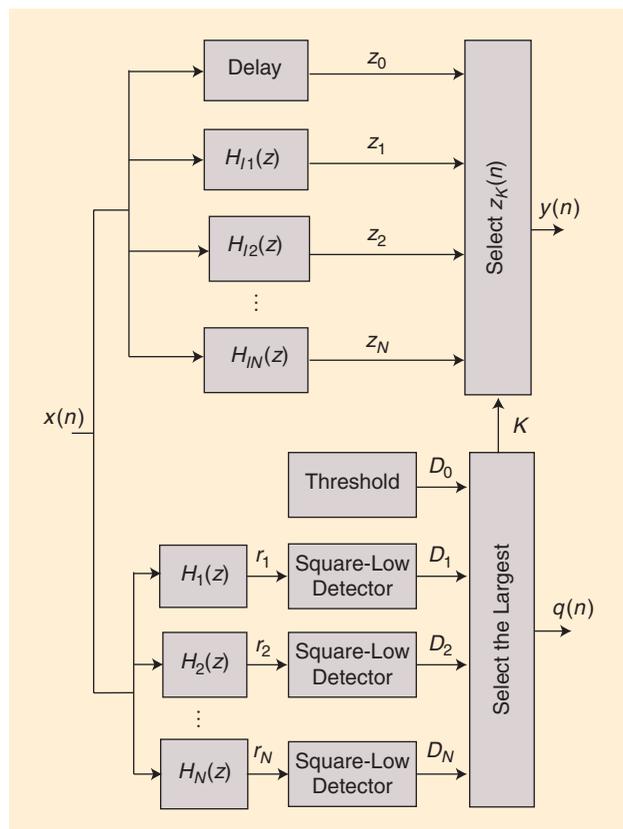


FIGURE 4. The resonance noise detection and removal system using resonant and inverse filter banks. (Figure adapted from [18].)

misidentification of cylinders could cause the control to diverge and negatively impact fuel economy and emissions.

Over the past two decades, many methods have been explored for determining the fuel-air ratio for individual cylinders. The majority use upstream oxygen sensors (O_2 sensors) located before the catalytic converters as the signal sources [22]–[32]. This type of O_2 sensors includes either switching or wide-range O_2 sensors. A few studies have focused on other methods such as ion current sensing [33].

An easy solution may be the direct measurement of each cylinder's fuel-air ratio by installing wide-range O_2 sensors inside each runner that is connected to its exhaust manifold. However, such a system is not practical or affordable for mass production. Thus, many methods to solve the problem have been researched and studied by automotive manufacturers and suppliers as well as academic institutes in recent years. They may be generally divided into two basic methods: model-based estimation methods and signal processing-based detection or estimation methods.

In the model-based methods [22]–[26], a physical, mathematical, or physical and mathematical mixed model is developed to describe the dynamic gas flows of an engine exhaust system. The model is usually considered as a state-space system with input of the O_2 signals measured at the exhaust system's confluence point. Then, Kalman filtering is applied to the state-space system to estimate the individual cylinder fuel-air ratio information. The estimation performance highly depends

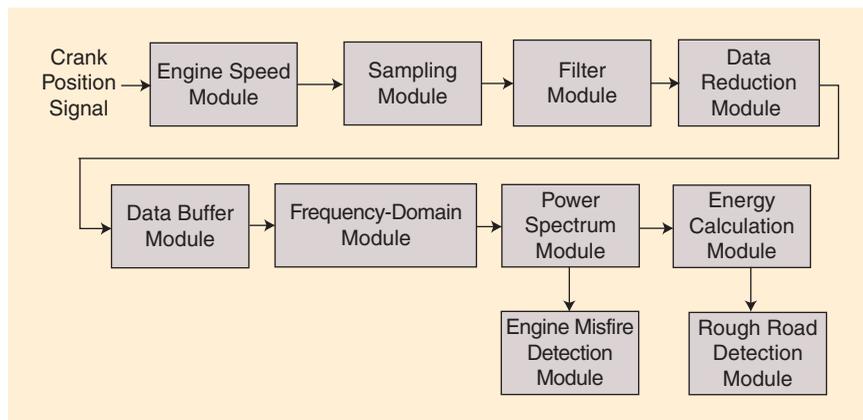


FIGURE 5. An FFT-based rough road detector. (Figure adapted from [20].)

on the developed model's accuracy and robustness against real-world noise and interference. In real applications with various engine operating conditions, the exhaust gas flows are very complex and usually nonlinear and time varying. Thus, not only the model development but also the determination of the model parameters is relatively complicated.

To reduce the complexity of the model and simplify the processing associated with updating estimation, some researchers have proposed linear models for individual cylinder fuel-air ratio estimation and application of recursive least-squares (RLS) techniques for the model parameter identification. In the method shown in [28], an engine and its exhaust system are modeled as an FIR system whose time-varying coefficients are determined based on measured engine sensor signals.

Signal processing methods, however, do not rely on any model and take advantage of the unique frequency characteristics of the O_2 signals when fuel-air ratio imbalance is present, thereby facilitating the retrieval of fuel-air ratio imbalance information [29]–[31]. The key mechanism used is that the upstream O_2 signal, either from a wide-range or switching O_2 sensor, will become rough when the imbalance occurs and the roughness is proportional to the severity of the fuel-air ratio imbalance. If the O_2 signal is sampled in an engine crankshaft angular domain, then the frequency with respect to crankshaft angular is constant and depends on the cylinder numbers in an engine exhaust bank. Thus, the imbalance signature can be easily retrieved using frequency filtering or FFT analysis techniques.

In [29], the O_2 signal is filtered using a high-pass IIR filter to obtain the fuel-air ratio imbalance frequency components. The fil-

tered signal is then separated into multiple signal sequences with an assumption that there is no gas coupling between cylinders. An open-loop fuel perturbation and signal sequence comparison procedure is used to select the signal sequence that is the most aligned to a cylinder. The cylinder alignment becomes necessary for the method because each cylinder-related signature in the filtered signal has variable delays due to gas flow characteristics in the exhaust system.

Figure 6 shows a filtering-based fuel-air ratio imbalance detector, in which a high-order FIR bandpass filter is used to retrieve the fuel-air ratio imbalance

frequency components [30]. The FIR filter order is 50 for a ten-cylinder engine. The filter used for the O_2 signal is designed in terms of an order-based spectral analysis and filtering technique. The bandpass-filtered signal is compared to a precalibrated threshold that depends on the engine speed and load for the fuel-air ratio imbalance decision making.

To further determine the fuel-air ratio imbalanced cylinder, the data that are output from the bandpass filter are sequentially and periodically saved into data buffers corresponding to the different engine cylinders. The buffer size is decided by the cylinder number in an engine exhaust bank. Then, each buffer's signal is low-pass filtered. The averaged signal in each buffer is counted in a given air-fuel ratio imbalance test window. Air-fuel ratio imbalance cylinder identification is determined only at the end of the test window. The cylinder ID associated with air-fuel ratio imbalance is identified as the one corresponding to the index of the imbalance counter that has a maximum counter value if the counter value is greater than the imbalance emission threshold. The index of the imbalance counter is calculated from the buffer index with an adjustment of exhaust gas transport delay cylinder event count offsets. Similar to the aforementioned method presented in [29], the accuracy of the exhaust gas transport delay estimation is also critical for the cylinder identification. In this method, the exhaust gas transport delay cylinder event count is calibrated based on vehicle dynamometer data for each engine operating point, engine speed and load, and both rich and lean imbalance conditions.

Another method is based on FFT or DFT analysis of O_2 signals for individual cylinder fuel-air ratio imbalance detection

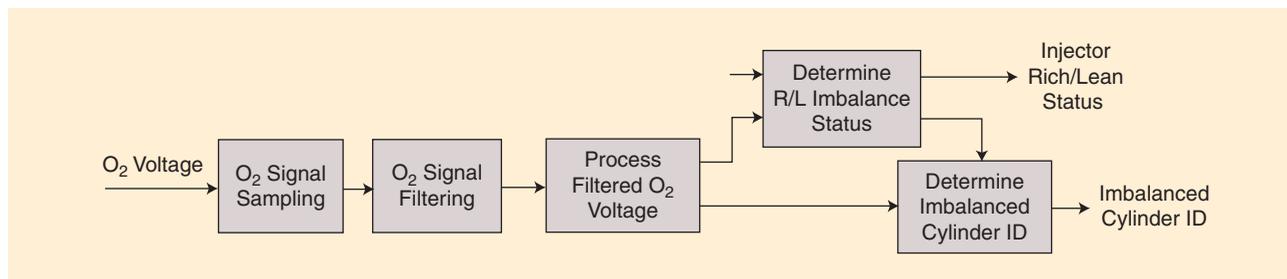


FIGURE 6. A filtering-based fuel-air ratio imbalance detector. (Figure adapted from [30]. Reprinted with permission from SAE International.)

and control [31]. In this method, the FFT or DFT of the O_2 signal is calculated first, and the amplitude of the first harmonic is then used for fuel-air ratio imbalance decision making through a comparison with a given threshold. Once the imbalance is detected, it retrieves fuel-air imbalance relations among cylinders through the study of the frequency characteristics of the cylinder imbalance patterns. With this method, no synchronization signal is needed, and thus it avoids the risks associated with possible synchronization errors.

The aforementioned methods are frequency-domain signal processing methods. A nonfrequency-domain-based signal processing method is presented in [32]. In the previously discussed methods, the obtained fuel-air ratio information is a relative quantity, not the estimate of the actual fuel-air ratio or its deviation from the stoichiometric point. The method shown in [32] uses a temporal array signal processing technique to estimate the actual fuel-air ratio or equivalently its deviation to the stoichiometric point, through an array of data samples from an oxygen sensor located in a confluence point of runners fed into each cylinder's estimator, as shown in the block diagram of Figure 7 [32]. For a switching O_2 sensor, in the method, the O_2 signal sampled in the crankshaft angular domain is first linearized, and then input into the signal array formation block to generate a signal array for the input of the fuel richness estimator, which is used to estimate the absolute fuel-air ratio or its equivalent. Finally, the estimated fuel-air ratio of each cylinder is used for OBD or fuel-air ratio balance control. This method requires the installation of wide-range sensors in each exhaust runner for the calibration of the fuel richness estimators.

In summary, model-based methods typically originated from control theoretical concepts. They are relatively complicated in design and calibration for mass production. Signal processing methods (and the frequency filtering methods, in particular) are more attractive to automotive manufacturers because of their implementation simplicity and relatively lower computational load and calibration efforts. Most filter parameters can be determined in the system development stage, instead of being calibrated in the vehicle release stage. This is a strong advantage for application of signal processing techniques.

Knock detection

Minimizing engine exhaust emissions and improving fuel economy are two challenges in modern engine controls. One approach to improving fuel economy is to maximize engine efficiency by optimally controlling spark timing under given operating conditions [34]. Abnormal combustion, such as combustion that occurs too early, will result in a knocking phenomenon that not only limits engine power generation but also causes damage to engine pistons, rings, exhaust valves, and other engine components. Thus, knock detection is crucial for optimal spark timing and engine protection.

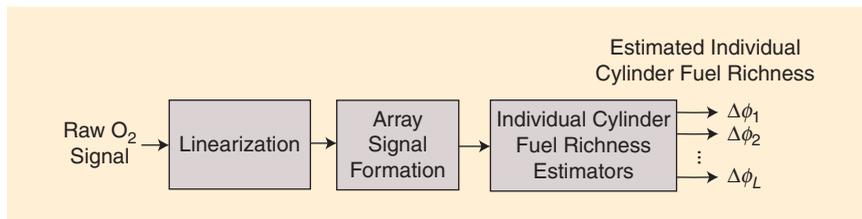


FIGURE 7. The system block for individual cylinder fuel richness estimation. (Figure adapted from [32]. Reprinted with permission from SAE International.)

When engine knocking occurs, pressure shock waves are generated and excite resonance inside the combustion chamber. The knock resonance frequency depends on the cylinder bore diameter and combustion chamber gas temperature [40] and typically ranges from 5 to 7 kHz. Two to four additional frequency peaks are also evident between the fundamental frequency and 20 kHz. Because of the characteristics of the knock signals, signal processing techniques are the most effective means to enhance the SNR and reliably perform knock detection.

Several types of sensors, such as the piezoelectric accelerometer sensor, cylinder pressure sensor, and ion sensor have been investigated for knock detection. Due to reasons of cost and technology maturity, however, the sensor most commonly used in production is the piezoelectric accelerometer knock sensor, which is usually mounted on the engine block and generates vibration signals. Regardless of sensor type, the signals generated are inevitably contaminated by noise or interference; thus, it is necessary to further process the raw signals to enhance the SNR to achieve reliable knock detection. Additionally, the signals measured from knock-generated shock waves and vibrations are generally nonlinear and time varying. This makes knock detection more complicated and challenging. Not only have automotive engineers been making significant efforts to develop practical and reliable knock detection systems, but many academic researchers have also been exploring advanced signal processing techniques for knock detection. The result has been that knock detection has become a very fruitful area for application of signal processing techniques in engine control and diagnostics.

Various knock detection methods that use signal processing have been intensively studied in recent decades. The methods mainly use frequency-domain filtering, DFT [38], [39], power spectrum analysis [40], time-frequency analysis [41]–[43], and wavelet transform [44]–[46]; refer to [35] for reviews of the early generation of knock detection algorithms used by the automotive industry.

Knock detection based on frequency-domain filtering has been widely used in automotive production because of its simplicity and low cost. Even today, it is still the main approach used in mass production automotive applications. As shown in Figure 8, typically, filtering methods use bandpass filters to retrieve the knock-based resonance signatures after preprocessing of the signal from a knock sensor, such as antialiasing filtering and signal amplifying, use an envelope detector (or rectifier) to obtain the knock intensity signal and then make a

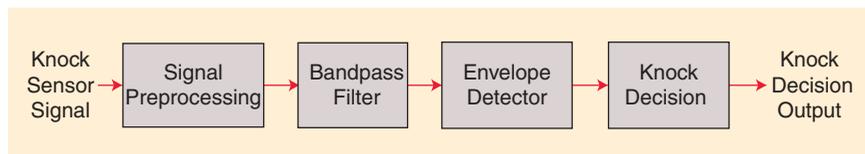


FIGURE 8. A block diagram of a typical frequency-domain-filtering-based knock detector.

knock decision. The implementation of filtering and envelope detection approaches for knock intensity detection evolved in several stages from analog technologies, to mixed analog and digital technologies, to today's fully digital technologies.

As mentioned previously, the knock signals' resonance frequency and its harmonics range from 5 to 20 kHz, and the signal sampling rate can be as large as 100 kHz. Existing common ECUs have difficulty performing real-time processing of signals with such a high sampling rate. Thus, knock detection is often implemented in a dedicated integrated circuit, and then the knock intensity is transferred from a chip to an ECU through a communication interface at a much slower data rate compared to the knock sampling rate. Today's knock detector devices have flexible configurations with programmable capabilities for gain, bandpass filter frequencies, and integrators.

When an ECU receives knock intensity signals from a knock detector device, the knock decision is made by comparing the knock intensity or its integration result to a preset threshold or the one calculated or obtained using look-up tables based on engine operating conditions, such as engine speed and load. To improve the detection performance and reduce calibration effort, some recent papers have proposed further processing of knock intensity using statistical methods. These stochastic approaches assume the knock intensity has a lognormal distribution and estimate a new metric, called the *knock factor*, that is related to the high and low percentiles for a lognormal distribution given the number of consecutive knock intensities [37]. The knock factor then is used for knock control.

In addition to frequency-domain-filtering methods, methods that perform spectrum analysis, such as power spectrum, FFT, and DFT, have advantages for knock detection [38]–[40]. Among them, however, DFTs are the most viable solution for real-time knock detection, at least until ECUs become computationally more powerful. A detailed strategy using DFTs is discussed in [38] for implementing the detection of knock signatures. The method uses multiple single-point DFTs to monitor the fundamental frequency plus the vibrational modes of an engine. The DFT algorithm provides better frequency discrimination than analog or digital filters. It is also less computationally intensive than an FFT when only a few frequency points are monitored, even though the FFT is more efficient in total time required when calculating across all allowable frequency ranges. In real-time processing, DFT has another advantage over FFT. That is, all samples for an FFT must be stored in memory before its calculations, but the DFT can be calculated one sample at a time because there is no linkage between samples as there is with the FFT [38]. Similar to time-based filtering, the DFT method uses a time window to process the data points in the knock-possible period. The paper

[39] presented an approach for varying the knock detection window length while using the single-point DFT detection method.

The aforementioned frequency filtering and transform spectrum methods are relatively simple. However,

because they do not consider the time-varying nature of resonance frequencies as well as nonlinear behavior of knock waves due to the complicated changes of cylinder gas temperatures and other physical phenomena, the knock information cannot be precisely detected or estimated for robust knock control and optimal spark timing under some engine operating conditions. To more effectively retrieve time-frequency dependent information associated with engine knocking, more advanced signal processing techniques such as time-frequency analyses and wavelet transforms have been studied recently [41]–[45].

The advantage of using time-frequency signal representations is in their ability to recognize time-frequency-dependent features, such as the frequency shifts that occur in the knock signal [41]. Both the Wigner–Ville distribution and the cross Wigner–Ville distribution were studied for a V6 engine's knock signal analysis in [41], and the test results showed that the SNR improved significantly compared to the bandpass filtering method. To overcome the drawback of the large number of operations required to perform such transforms, a pseudo-Wigner distribution was proposed for the time-frequency analysis of the knock signals [42]. Figure 9 shows a typical high-pass filtered knock signal and its Wigner–Ville spectrum generated from several knock signals, where the knock resonance frequencies are clearly displayed with changes corresponding to the engine crank angle [48].

The wavelet transform is good for time-scale analysis of a signal. In recent years, many researchers, mainly from academic areas, studied the application of wavelet transforms for knock feature detection. Unlike the Fourier transform, the wavelet transform provides the time-evolution of the signal at different scales. The discrete wavelet transform (DWT) is a computationally efficient implementation of the wavelet transform and is more suitable for knock detection applications as reported by [44]. The paper [45] proposes a knock feature extraction method using wavelet packet transforms, and through the examples they studied, they showed that the wavelet packet transform improves the time-frequency resolution and has advantages in extracting knock feature information, knock intensity and the time of knock occurrences, even under light knock conditions, compared to the DWT method.

Instead of using accelerometer sensor signals for knock analysis, some researchers have also studied wavelet transforms using ion current signals [46].

In addition to the aforementioned mainstream knock detection techniques, other signal processing methods were also explored for knock detection recently, such as neural networks-based knock feature extraction methods [47], model-based knock detection methods [48], as well as the Hilbert–Huang

transform (HHT)-based knock detection method [49], which allows for the decomposition of a multicomponent nonstationary signal into individual components. A precise characterization of the knock phenomenon can be observed from separated knock resonance components using HHT.

Signal processing approaches for knock detection have become a very productive approach to supporting engine control and diagnostics. Due to today's ECU processing capabilities, the most popular methods used in mass production applications are still those based on filtering. DFT methods are becoming feasible for emerging ECUs with high-performance processors or with additional dedicated DSP coprocessors or special hardware systems. It is still not practical for many advanced signal processing methods, such as time-frequency analysis and wavelet transform techniques, to be performed on an ECU in real time; however, these techniques can possibly be used in off-board analyses during production development stages.

Summary and outlook

This article has provided an overview of how DSP is being applied in several important areas associated with engine control and OBD. At a high level, the DSP techniques used in these applications mainly involve frequency filtering, FFTs and DFTs, time-frequency signal analysis, and wavelet transforms. Additionally, a wide array of advanced signal processing techniques, including Kalman filtering, neural networks, the HHT, and other techniques, have been studied and shown to be promising, and the industry is anticipating that these methods will become more valuable as their efficiency increases and their cost of implementation decreases in the future. DSP techniques show great promise for solving many challenging problems in misfire detection, estimating individual cylinder fuel-air ratios, and performing knock detection. Some of the DSP algorithms that have been developed demonstrate unique advantages that are difficult to obtain using other non-DSP methods, particularly in terms of achieving high accuracy and robustness when performing event detection and estimation of underlying parameters.

As one surveys these different DSP applications for automobiles, what is evident is that over the past few decades there has been an evolution in how DSP techniques have been researched and applied to engine control and diagnostics. This evolution is basically a reflection of the challenges facing automobile powertrain applications before they can be successfully integrated into a production application: a newly developed method is required to address an important safety and regulatory requirement, while achieving robust performance, being able to carry out its processing in real time, having low implementation cost, and requiring minimal to low calibration effort.

These challenges are complex, especially as an engine needs to run in various ambient environments and under various specified operating conditions. Unfortunately, many methods can work for a special case or in a limited range, but they are still not useful for real automotive applications because they are unable to fulfill all of the required tasks faced in a

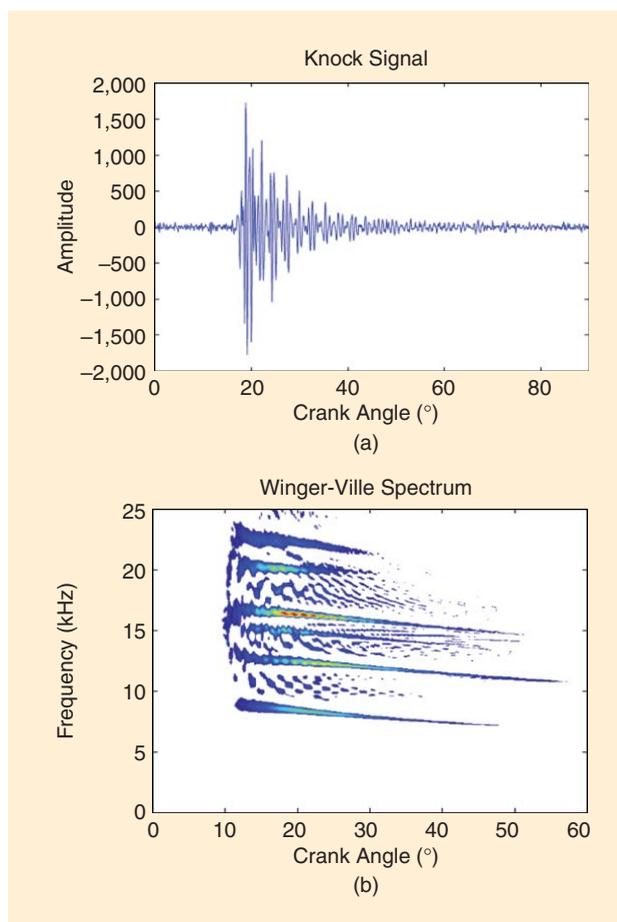


FIGURE 9. (a) High-pass filtered knock pressure signal; (b) Wigner-Ville spectrum of several knock signals. (Figure reproduced from [48].)

real production vehicle. Thus, a useful DSP method has to be robust enough to handle all these situations in production without any exception.

Because of real-time engine control and OBD, DSP methods also should not be too complicated so that they can be processed within a permitted time slot, while all of the other processing tasks are being performed. For example, the time interval between two engine firings is only 2.5 ms for an eight-cylinder engine at an engine speed of 6,000 rev/min. Frequency-domain-filtering techniques have a relatively low computational time compared to most advanced DSP techniques. Since many engine signals have periodic characteristics, automotive engineers have artfully designed DSP algorithms that take full advantage of these characteristics to make their system design and sampling strategies more robust and simpler, without having to rely blindly on heavy-hammer approaches to solving problems.

Although ECU processing power continues to increase, the engine control and diagnostic algorithms become more and more sophisticated and complex thus outpacing available computational resources. The application of many advanced algorithms is currently limited by ECU processing capability. More powerful processing capability means more costs

for production. Cost is an important consideration for automotive manufacturers. Low cost and high processing performance of the ECUs are needed for mass production. Beyond the intensive research on various efficient and robust DSP algorithms, microprocessor manufacturers have focused their efforts on building computationally powerful automotive oriented microcontrollers with the goal of reducing overall costs by utilizing various technologies. For example, the Motorola CPU16 has some very basic DSP instructions, and Freescale Semiconductor's MPC 5554 Copperhead microcontroller and MPC5674F Power Architecture microcontroller have a powerful built-in signal processing engine (SPE), which is able to accelerate signal processing with its vector processing capability that allows two arithmetic operations to be completed in parallel. The SPE works with either fixed-point or floating-point operations and runs especially efficiently for DSP algorithms because the specially designed DSP functionality can accelerate signal processing computations, such as those associated with FIR and IIR filters and DFTs.

According to a study of software solutions for knock detection by [50], more than 50% reduction of ECU computational time can be achieved using the Copperhead microcontroller's SPE functions compared to the conventional C code implementation, and, therefore, the use of the SPE feature makes an ECU capable of running more applications or extending the life of one generation of ECU. The microcontroller MPC5674F further provides a special feature to perform frequency-filtering-based knock detection without additional external components. This avoids use of separate knock detection application-specific integrated circuits and other components and therefore reduces overall system cost. The aforementioned example represents a trend of comprehensive on-chip integration and virtual sensing capabilities provided in high-performance automotive microcontrollers to help automotive developers accelerate next-generation engine control design and simultaneously achieve both performance and cost benefits.

From the viewpoint of automotive manufacturers, there are many vehicle models to be produced every year. For each feature or application, they wish to have a unified algorithm that can be used in all models of vehicles through proper configuration and minimal calibration complexity. The development of DSP applications needs to consider these requirements. Because of the periodic characteristics of engine signals, DSP has the potential to reduce calibration effort. Some interested frequencies are constant when the systems are designed to execute in the crankshaft angle domain and use signals that are sampled in the same domain. This significantly simplifies the designed DSP systems and reduces the calibration efforts. As in the previously given examples of misfire detection and individual cylinder fuel-air ratio imbalance detection, frequency-domain DSP methods have a distinct advantage.

Based on the aforementioned criteria and considerations, it is not difficult to understand why the most popular DSP methods for engine control and OBD are still frequency-filtering techniques—particularly because of the techniques'

maturity and relatively low implementation cost. The FFT or DFT methods are becoming feasible when they are implemented in the ECUs with high-performance processors or ones with additional dedicated DSP features or special DSP hardware systems. It is still not practical for many advanced signal processing methods, such as time-frequency analysis and wavelet transforms, to be deployed for onboard real-time processing in today's ECU environment because of their high computational demands. However, these techniques can possibly be used in off-board analyses during production development stages. We view these challenges as a great opportunity for signal processing to play a critical role in the automotive industry. As the title of a January 2014 "Special Reports" column in *IEEE Signal Processing Magazine* [51] indicated, the automotive industry is a key contributor to the success of the DSP sector. There are many opportunities for DSP applications in automobiles, including infotainment, telematics, advanced driver assistance systems, and autonomous driving. DSP applications in powertrain controls seem not as visible to most automobile consumers as the aforementioned application areas for making smarter cars. However, one trend is evident: advanced DSP techniques will be essential in various engine control and OBD applications, as well as in other powertrain controls, especially as ECUs continue to evolve to have more processing power available for more mature and sophisticated DSP algorithms that will improve vehicle fuel economy and emission performance for making greener cars.

Acknowledgments

We would like to thank Dr. Wade Trappe for his many valuable editing suggestions that were adopted in the manuscript.

Authors

Zhijian James Wu (zhijian.wu@gm.com) received his M.S.E. and Ph.D. degrees from the University of Michigan, Ann Arbor, in 1990 and 1991, respectively, and holds bachelor's and master's of engineering degrees in electronic engineering from Shanghai Jiao Tong University, China (1982 and 1984, respectively). He has more than 20 years of R&D experience in engine controls and onboard diagnostics. He joined General Motors in 2014. He was previously a senior technical specialist with Fiat Chrysler Automobiles. He has more than 20 patents and technical publications and is a Senior Member of the IEEE and a member of the Society of Automotive Engineers International.

Sanjeev M. Naik (sanjeev.m.naik@gm.com) holds a bachelor's degree from IIT Bombay (1986), an M.S.E.E. degree from the University of Michigan, Ann Arbor (1988), and a Ph.D. degree from the University of Illinois, Urbana-Champaign (1992), all in electrical engineering, and an M.B.A. degree from the University of Michigan, Ann Arbor (2000). He is the engineering group manager for controls advanced engineering at General Motors (GM). Prior to GM, he worked at Cummins Engine Company. He has several publications and 50 patents. His technical interests are in the

application of advanced control and signal processing techniques to automotive problems. He is a Senior Member of the IEEE and a member of SAE International.

References

- [1] S. Chen, L. Chien, M. Nagashima, J. Van Ess, and S. Hashemi, "Misfire detection in a dynamic skip fire engine," *SAE Int. J. Engines*, vol. 8, no. 2, pp. 389–398, 2015.
- [2] M. Klenk, W. Moser, W. Mueller, and W. Wimmer, "Misfire detection by evaluating crankshaft speed: A means to comply with OBDII," SAE Tech. Paper 930399, 1993, doi: 10.4271/930399.
- [3] Y. Wang, K. Huang, Z. Lin, and F. Meng, "Advanced gasoline engine misfire diagnostic method based on crankshaft speed multiple filtering," in *Int. Conf. Electric Information Control Engineering*, 2011, pp. 1964–1968.
- [4] Z. Wu and A. Lee, "Real-time misfire detection for automobile engines with medium data rate crankshaft sampling," U.S. Patent 6 006 155A, Dec. 1999.
- [5] T. Aono and E. Fukuchi, "Misfire Detection method robust against crankshaft vibration and acceleration," in *Proc. IEEE Conf. Control Applications*, Toronto, Canada, Aug. 28–31, 2005, pp. 1218–1221.
- [6] X. Ma, Z. Xia, H. Wu, and X. Huang, "Combined frequency domain analysis and fuzzy logic for engine misfire diagnosis," SAE Tech. Paper 2015-01-0207, 2015, doi: 10.4271/2015-01-0207.
- [7] X. Yu, L. Liu, S. Yang, X. Zhong and D. Liu, "Misfire detection based on engine speed using wavelet," in *Proc. IEEE Intelligent Vehicles Symp.*, 2009, pp. 833–837.
- [8] M. Montini and N. Speciale, "Multiple misfire identification by a wavelet-based analysis of crankshaft speed fluctuation," in *Proc. IEEE Int. Symp. Signal Processing Information Technology*, 2006, pp. 144–148.
- [9] S. Naik, "Advanced misfire detection using adaptive signal processing," *Int. J. Adapt. Control Signal Process. Special Issue Automat. Appl.*, vol. 18, no. 2, pp. 181–198, Mar. 2004.
- [10] A. Lee, R. N. K. Loh, and Z. Wu, "Automotive engine misfire detection using Kalman filtering," in *Proc. 58th Vehicular Technology Conf.*, 2003, vol. 5, pp. 3377–3381.
- [11] S. Helm, M. Kozek, and S. Jakubek, "Combustion torque estimation and misfire detection for calibration of combustion engines by parametric Kalman filtering," *IEEE Trans. Ind. Electron.*, vol. 59, no. 11, pp. 4326–4337, 2012.
- [12] Z. Wu and A. Lee, "Misfire detection using a dynamic neural network with output feedback," SAE Tech. Paper 980515, 1998, doi: 10.4271/980515.
- [13] N. Cavina, L. Poggio, and G. Sartoni, "Misfire and partial burn detection based on ion current measurement," *SAE Int. J. Engines*, vol. 4, no. 2, pp. 2451–2460, 2011.
- [14] K. Doi, Y. Nakamura, K. Hanashi, and K. Hashizume, "Development of spark plug for ion current misfire detection system," *SAE Int. J. Engines*, vol. 5, no. 3, pp. 1387–1393, 2012.
- [15] J. Williams, "An overview of misfiring cylinder engine diagnostic techniques based on crankshaft angular velocity measurements," SAE Tech. Paper 960039, 1996, doi: 10.4271/960039.
- [16] K. Iida, K. Akishino, and K. Kido, "IMEP estimation from instantaneous crankshaft torque variation," SAE Tech. Paper 900617, 1990, doi: 10.4271/900617.
- [17] K. Nishida, T. Kaneko, Y. Takahashi, and K. Aoki, "Estimation of indicated mean effective pressure using crankshaft angular velocity variation," SAE Tech. Paper SAE 2011-32-0510, 2011, doi: 10.4271/2011-32-0510.
- [18] Z. Wu and A. Lee, "A resonance noise detection and removal technique using resonant and inverse filter banks and its application in engine misfire detection," in *Proc. IEEE EIT Electro Information Technology Conf.*, Oakland University, June 6–8, 2001.
- [19] J. Asgari, T. Feldkamp, C. Stephens, M. Fridman, D. Hrovat, A. Varady, and M. McConnell, "Rough road detection system used in an on-board diagnostic system," U.S. Patent 8 060 275B2, Nov. 2011.
- [20] C. Hernandez, "Rough road detection system," U.S. Patent 7 299 687B2, Nov. 2007.
- [21] T. Assaf, D. Mathews, and S. Naik, "Signal processing for rough road detection," *SAE Int. J. Passeng. Cars Electron. Electr. Syst.*, vol. 3, no. 1, pp. 81–87, 2010.
- [22] Y. Hasegawa, S. Akazaki, I. Komoriya, H. Maki, Y. Nishimura, and T. Hirota, "Individual cylinder air-fuel ratio feedback control using an observer," SAE Tech. Paper 940376, 1994, doi: 10.4271/940376.
- [23] R. Turin and S. Naik, "Individual cylinder fuel control method," U.S. Patent 6 314 952B1, Nov. 2001.
- [24] J. Grizzle, K. Dobbins, and J. Cook, "Individual cylinder air-fuel ratio control with a single EGO sensor," *IEEE Trans. Veh. Technol.*, vol. 40, no. 1, pp. 280–286, 1991.
- [25] J. Chauvin, P. Moulin, G. Corde, N. Petit, and P. Rouchon, "Kalman filtering for real-time individual cylinder air fuel ratio observer on a diesel engine test bench," in *Proc. American Control Conf.*, Minneapolis, MN, 2006, pp. 1887–1891.
- [26] B. He, T. Shen, J. Kako, and M. Ouyang, "Input observer-based individual cylinder air-fuel ratio control: Modeling, design and validation," *IEEE Trans. Control Syst. Technol.*, vol. 16, no. 5, pp. 1057–1065, 2008.
- [27] Y. Liu, T. Shen, K. Sata, and K. Suzuki, "Modeling of individual cylinder air-fuel ratio for IC engines with multi-cylinders," in *Proc. 30th Chinese Control Conf.*, Yantai, China, 2011, pp. 6151–6156.
- [28] J. Meyer, S. Yurkovich, and S. Midlam-Mohler, "A model based estimator for cylinder specific air-to-fuel ratio corrections," *J. Dyn. Syst. Meas. Control*, vol. 1331, no. 3, pp. 031001-1–031001-14, Mar. 2011.
- [29] J. Kainz and J. Smith, "Individual cylinder fuel control with a switching oxygen sensor," SAE Tech. Paper 1999-01-0546, 1999, doi: 10.4271/1999-01-0546.
- [30] N. Qiao, C. Krishnamurthy, and N. Moore, "Determine air-fuel ratio imbalance cylinder identification with an oxygen sensor," *SAE Int. J. Engines*, vol. 8, no. 3, pp. 1011–1005, 2015.
- [31] H. Javaherian, A. Brown, and M. Nolan, "Individual cylinder air-fuel ratio control part I: L3 and V6 engine applications," SAE Tech. Paper 2011-01-0695, 2011, doi: 10.4271/2011-01-0695.
- [32] Z. Wu and B. Wasacz, "Estimation of individual cylinder fuel-air ratios from a switching or wide range oxygen sensor for engine control and on-board diagnosis," *SAE Int. J. Engines*, vol. 4, no. 1, pp. 813–827, 2011.
- [33] N. Wickstrom, S. Byttner, and U. Holmberg, "Robust tuning of individual cylinders AFR in SI engines with the ion current," SAE Tech. Paper 2005-01-0020, 2005, doi: 10.4271/2005-01-0020.
- [34] E. Corti and C. Forte, "Real-time combustion phase optimization of a PFI gasoline engine," SAE Tech. Paper 2011-01-1415, 2011, doi: 10.4271/2011-01-1415.
- [35] O. Boubai, "Knock detection in automobile engines," *IEEE Instrum. Measure. Mag.*, vol. 3, no. 3, pp. 24–28, Sept. 2000.
- [36] J. Wagner, J. Keane, R. Koseluk, and W. Whitlock, "Engine knock detection: Products, tools, and emerging research," SAE Tech. Paper 980522, 1998, doi: 10.4271/980522.
- [37] Y. Lonari, C. Polonowski, J. Naber, and B. Chen, "Stochastic knock detection model for spark ignited engines," SAE Tech. Paper 2011-01-1421, 2011, doi: 10.4271/2011-01-1421.
- [38] T. Horner, "Knock detection using spectral analysis techniques on a Texas instruments TMS320 DSP," SAE Tech. Paper 960614, 1996, doi: 10.4271/960614.
- [39] M. Kearney, "Knock signal conditioning using the discrete Fourier transform and variable detection window length," SAE Tech. Paper 2007-01-1509, 2007, doi: 10.4271/2007-01-1509.
- [40] D. Lazarescu, V. Lazarescu, and M. Ungureanu, "Knock detection based on power spectrum analysis," in *Proc. Int. Symp. Signals, Circuits and Systems*, 2005, vol. 2, pp. 701–704.
- [41] B. Samimy, G. Rizzoni, and K. Leisenring, "Improved knock detection by advanced signal processing," SAE Tech. Paper 950845, 1995, doi: 10.4271/950845.
- [42] S. Ker, F. Bonnardot, and L. Duval, "Algorithm comparison for real time knock detection," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2007, vol. 2, pp. 397–400.
- [43] G. Matz and F. Halawatsch, "Time-frequency methods for signal detection with application to the detection of knock in car engines," in *Proc. IEEE Statistical Signal and Array Processing*, 1998, pp. 196–199.
- [44] J. Borg, G. Saikalas, S. Oho, and K. Cheok, "Knock signal analysis using the discrete wavelet transform," SAE Tech. Paper 2006-01-0226, 2006, doi: 10.4271/2006-01-0226.
- [45] C. Liu, Q. Gao, Y. Jin, and W. Yang, "Application of wavelet packet transform in the knock detection of gasoline engines," in *Proc. Int. Conf. Image Analysis and Signal Processing*, 2010, pp. 686–690.
- [46] Z. Zhang, N. Saiki, H. Toda, T. Imamura, and T. Miyake, "Detection of knocking by wavelet transform using ion current," in *4th Int. Conf. Innovative Computing, Information and Control*, 2009, pp. 1566–1569.
- [47] K. Kamal, M. Farid, and S. Mathavan, "Detection of knocking in spark ignition (SI) engines using CMAC neural networks," in *Proc. Int. Conf. Robotics and Artificial Intelligence*, 2012, pp. 144–148.
- [48] M. Urlaub and J. F. Bohme, "Evaluation of knock begin in spark ignition engines by least squares," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2005, vol. 5, pp. 681–684.
- [49] J. Fiolka, "Application of Hilbert-Huang transform to engine knock detection," in *Proc. 20th Int. Conf. Mixed Design Integrated Circuits and Systems*, 2013, pp. 457–461.
- [50] Z. Wu and L. Chen, "Method to efficiently implement automotive application algorithms using signal processing engine (SPE) of copperhead microcontroller," SAE Tech. Paper 2008-01-1222, 2008, doi: 10.4271/2008-01-1222.
- [51] R. Schneiderman, "Automotive industry is a key component to the success of the DSP sector," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 18–21, Jan. 2014.

Bashar I. Ahmad, James K. Murphy, Simon J. Godsill,
Patrick M. Langdon, and Robert W. Hardy

Intelligent Interactive Displays in Vehicles with Intent Prediction

A Bayesian framework

Using an in-vehicle interactive display, such as a touch screen, typically entails undertaking a freehand pointing gesture and dedicating a considerable amount of attention, that can be otherwise available for driving, with potential safety implications. Due to road and driving conditions, the user's input can also be subject to high levels of perturbations resulting in erroneous selections. In this article, we give an overview of the novel concept of an intelligent predictive display in

vehicles. It can infer, notably early in the pointing task and with high confidence, the item the user intends to select on the display from the tracked freehand pointing gesture and possibly other available sensory data.

Accordingly, it simplifies and expedites the target acquisition (pointing and selection), thereby substantially reducing the time and effort required to interact with an in-vehicle display. As well as briefly addressing the various signal processing and human factor challenges posed by predictive displays in the automotive environment, the fundamental problem of intent inference is discussed, and a Bayesian formulation is introduced. Empirical evidence from data collected in instrumented cars is shown to demonstrate the usefulness and effectiveness of this solution.

Introduction

The complexity of in-vehicle infotainment systems (IVIS) has been steadily increasing to accommodate the growing additional services associated with the proliferation of smart technologies in modern vehicles. They aim to improve the driving experience and safety, for example, advanced driver assistance, route guidance, driver inattention monitoring, and many others [1]. Consequently, minimizing the effort and distraction of interacting with or controlling the IVIS is a key challenge [2]. This article introduces and presents an overview of the predictive in-vehicle display system, which utilizes suitable statistical signal processing algorithms to enhance and simplify human-machine interaction (HMI) in automotive applications, including IVIS-related interactions.



ISTOCKPHOTO.COM/CHIBOSAN

Digital Object Identifier 10.1109/MSP.2016.2638699
Date of publication: 3 March 2017

Lately, there has been a strong move toward replacing traditional static mechanical controls in vehicles, such as buttons, switches, and gauges, with interactive displays, mainly touch screens [2]. This is motivated by the evolution of the increasingly ubiquitous touch-screen technology and the ability of these displays to

- 1) effectively handle a multitude of functions by incorporating large quantities of information associated with IVIS
- 2) promote intuitive interactions via freehand pointing gestures, especially for novice users
- 3) offer design flexibility through a combined display-input-feedback module
- 4) minimize clutter in the vehicle interior given their adaptability to the context of use, unlike mechanical controls.

For example, the Tesla Model S car features a 17-in touch screen controlling most of the car functions [3]. Additionally, other types of displays, such as head-up displays (HUDs) and general three-dimensional/dimensions (3-D) displays, have the potential of providing a more immersive driving experience and are becoming increasingly commonplace in vehicles [4], [5], for instance, the Jaguar Land Rover HUD windscreen incorporating laser holography [6]. However, such displays are often passive, and users lack the means to easily interact with them in an automotive setting.

Interacting with an in-vehicle touch screen typically involves undertaking a freehand pointing gesture to select an item on the display. This requires dedicating a considerable amount of visual, cognitive, and manual attention that is otherwise available for driving. The user input can also be subject to perturbations due to the road and driving conditions, resulting in incorrect on-screen selections [7], [8]. For example, the rate of successfully selecting an icon on the in-car display can be less than 50% when driving over a badly maintained road [8]. Rectifying an erroneous selection or adapting to the present noise will tie up more of the user's attention. This can render using interactive displays in vehicles effortful and distracting, with potential safety consequences [9]. Hence, there is a need for a solution that simplifies interaction with in-vehicle displays via intuitive freehand pointing gestures or that even enables it for emerging display technologies such as HUDs.

An intelligent in-vehicle predictive display, whose top-level block diagram is depicted in Figure 1, employs a gesture tracker (and possibly other sensory data when available), in conjunction with a probabilistic prediction algorithm to determine the item the user intends to select on the display, remarkably early in the freehand pointing gesture [10]. It subsequently facilitates and expedites the target acquisition. Thus, the introduced intent-aware

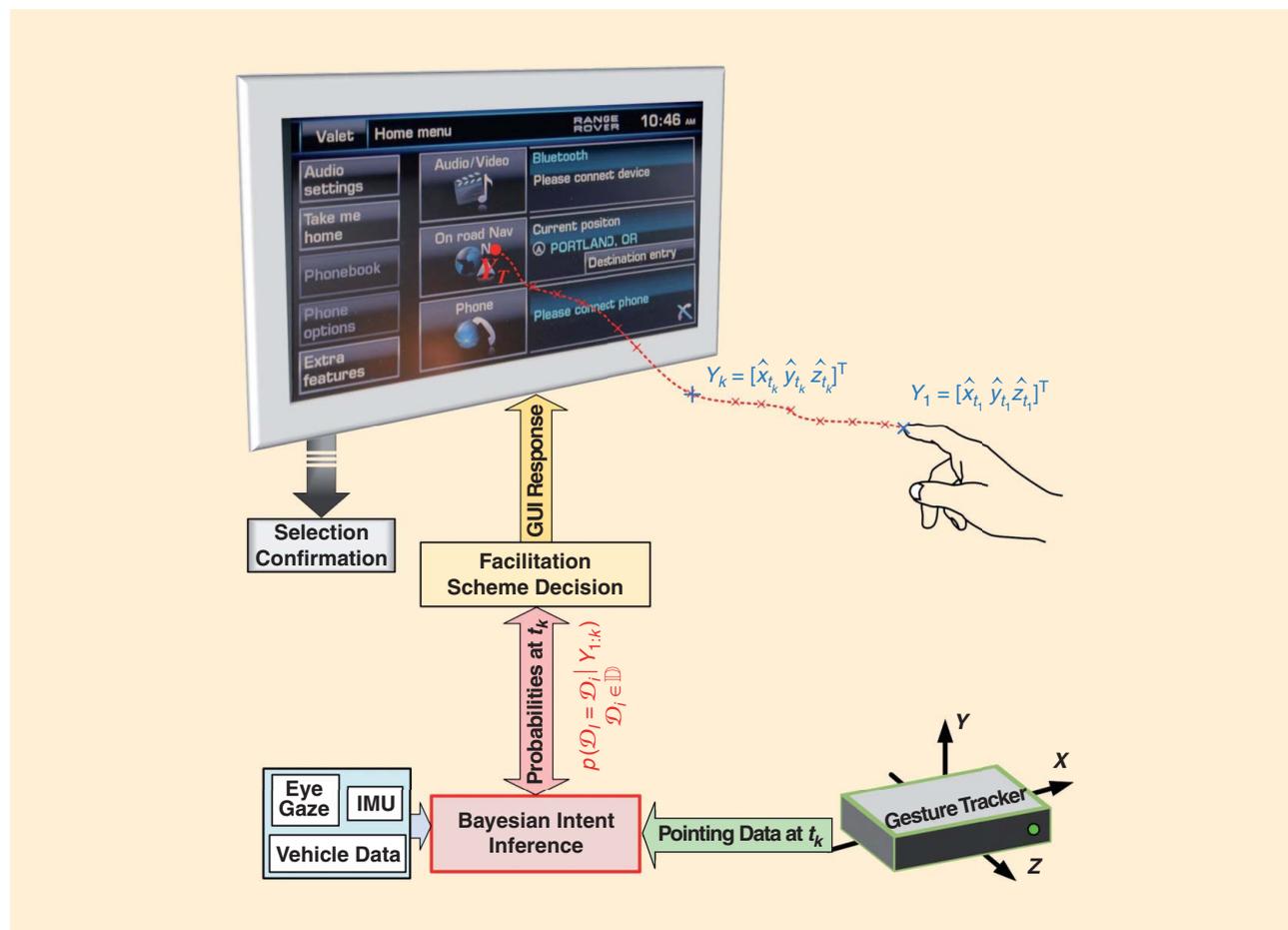


FIGURE 1. A block diagram of an in-vehicle predictive touch-screen system with a full 3-D pointing fingertip track, $t_k > t_1$. An infotainment menu of a Range Rover interface is displayed; the vehicle interior is not shown.

system can significantly improve the interactive display usability in vehicles and reduce the effort (attention) they require. Assuming that the prediction certainty meets a set criterion, the user need not touch the display surface to select the intended on-screen item, allowing midair selection. Therefore, this solution can also enable interaction with displays that do not have a physical surface, e.g., HUD and 3-D displays or projections.

This article highlights and gives a unified treatment of the various signal processing (e.g., tracking-filtering, fusion, prediction, etc.) and human factors (e.g., feedback, prior experience, etc.) challenges posed by the in-vehicle intent-aware display concept, some of which were individually considered in previous publications (including those for nonautomotive applications), such as [10]–[19]. In particular, the fundamental problem of intent inference within a Bayesian framework is addressed here, and suitable probabilistic prediction models are presented; they lead to a low-complexity implementation of the inference routine. Within this formulation, the task of smoothing perturbed pointing trajectories due to road and driving conditions via statistical filtering is discussed. The sensory requirements of the predictive system in the vehicle environment are also briefly outlined. Data collected in instrumented cars and results from a prototype predictive touch-screen system are shown to demonstrate the capabilities of this intelligent HMI solution.

Background

According to the renowned human movement model Fitts' law [20], the index of difficulty (ID) and total time (T) of

acquiring an interface icon (i.e., pointing and selection) are given by

$$ID = \log_2(1 + \ell/W),$$

$$T = a + b \log_2(1 + \ell/W), \quad (1)$$

where W and ℓ are the the width of the target item and its distance from the starting position of the pointing object (mouse cursor or pointing finger), respectively [12]; a and b are empirically estimated. As intuitively expected, the selection task can be simplified and expedited by applying a pointing facilitation scheme, such as increasing the item size (larger W) or moving it closer to the cursor (smaller ℓ). Since a typical graphical user interface (GUI) contains several selectable items, any assistive pointing strategy should be preceded by a predictor to identify the intended on-screen icon [12]. Hence, the end-point prediction problem has received notable attention in the human–computer interaction (HCI) area, e.g., [11]–[14] (see [10] and [14] for a brief overview).

The majority of existing HCI studies focus on pointing in two-dimensional/dimensions (2-D) via a mouse or mechanical device on a computer screen to acquire GUI icons. They often use deterministic pointing kinematics models for end-point prediction assuming 1) the pointing object (cursor) velocity has a consistent profile and is zero at arrival at destination, and 2) the cursor heads at a nearly constant angle toward its end point. Both premises make intuitive sense for mouse pointing in 2-D, however, they do not necessarily hold for freehand pointing gestures in 3-D [10]. For example, Figure 2(a) shows that the pointing fingertip heading angle to an on-screen icon drastically changes throughout a sample of freehand pointing gestures recorded in an instrumented car; d_I is the location of the intended on-screen destination in 3-D and Y_k is the 3-D Cartesian coordinates of the pointing fingertip at the time instant t_k .

Data-driven prediction techniques, such as in [13] and [19], can be applied to infer the intended destination of a pointing task. They often utilize a pointing motion model learnt from a priori recorded interactions, necessitating the availability of a complete data set of training examples of pointing trajectories. This requirement is particularly stringent for freehand gestures approaching a display in 3-D to select icons on GUIs of various possible layouts, due to the very large number of possible paths. Additionally, in an automotive HMI context, a user might be expected to only undertake a few pointing gestures, for instance, to set up the IVIS preferences, during his or her first system use, i.e., a very limited set of training tracks is often available. On the other hand, the predictive display system discussed here employs known motion as well as sensor models, and thus can use a state-space-modeling approach, albeit with a few unknown parameters. It requires minimal training and is computationally efficient.

In the area of object tracking, e.g., in surveillance applications, knowing the destination of a tracked object not only leads

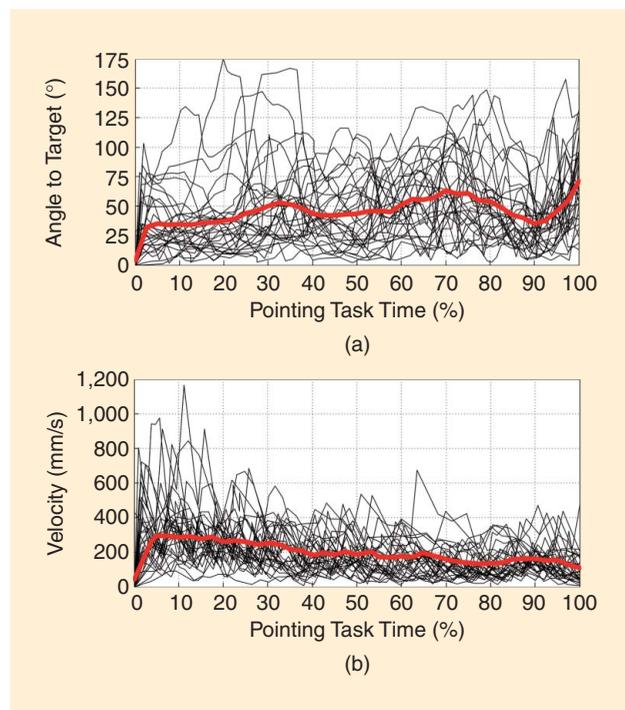


FIGURE 2. The angle to on-display icon and velocity profile for 30 in-car pointing tasks; the thick red line is the mean. (a) The angle to the intended GUI icon $\theta_k \triangleq \angle(Y_k - Y_{k-1}, d_I)$. (b) The pointing fingertip velocity $\|Y_k - Y_{k-1}\|_2$.

to more accurate tracking results, but also offers vital information on intent, revealing potential conflict or threat [16], [21], [22]. Destination prediction can be viewed as a means to assist planning and decision making at a system level higher than that of established conventional sensor-level tracking algorithms, whose objective is to infer the current value of the latent state X_t (e.g., the tracked object position, velocity, etc.) [22]. For example, destination-aware trackers that include an additional mechanism

to determine the object end point are proposed in [16]. These methods discretize the state space area into predefined regions, and the object can only pass through a finite number of these zones; such a discretization can be a burdensome task for free-hand pointing gestures in 3-D. On the contrary, the predictive display solution presented in this article uses continuous state space motion models that do not impose any restrictions on the path the pointing finger has to follow to reach its intended on-display end point and can easily handle noisy as well as asynchronous observations. Nevertheless, other conforming destination-aware tracking methods can be applicable.

A related scenario in which there is a growing interest is the user input on a smartphone, perturbed due to situational impairment, for example, walking [17]. Typically in such cases, the GUI is dynamically adapted to compensate for the measured noise. For an in-vehicle display, the pointing time and distance are notably longer than that for a handheld device and the correlation between the pointing hand movements and the experienced in-car accelerations or vibrations can be ambiguous [10]. This is attributed to the complexity of the human motor system and its response to noise as well as the seat position, cushioning, reaching style or distance, etc. Thus, compensating for the measured in-vehicle noise can have limited effects on improving the display usability. Here, perturbed user input is tackled within the statistical inference framework of a predictive display.

An in-vehicle predictive display system

Next, we describe the various modules that compliment the present in-vehicle interactive display, e.g., a touch screen, to realize the intelligent predictive display system in Figure 1.

Gesture tracker

Motivated by extending HCI beyond traditional keyboard input and mouse pointing, new 3-D vision sensory devices have emerged that can track, at high rates, hand gestures, including pointing fingertip(s), e.g., Microsoft Kinect, Leap Motion (LM), and SoftKinetic DepthSense. However, operating in a mobile vehicle environment can be challenging to these trackers due to dynamically changing light conditions, in-car vibrations-accelerations, occlusion with limited in-car mounting positions, large coverage area (e.g., steering wheel or armrest to display and the front passenger), and others. Fortunately, the current interest in gesture-based HCI in cars (e.g., current

Other types of displays, such as head-up displays and general three-dimensional displays, have the potential of providing a more immersive driving experience and are becoming increasingly commonplace in vehicles.

BMW 7 Series cars have gesture control for some features) is driving the development of automotive-grade gesture trackers [15]. In Figure 1, a tracker provides, in real time, the pointing hand/finger(s) locations, $Y_{1:k} \triangleq \{Y_1, Y_2, \dots, Y_k\}$, at the discrete time instants t_1, t_2, \dots, t_k . For instance, $Y_n = [\hat{x}_{t_n} \ \hat{y}_{t_n} \ \hat{z}_{t_n}]^T$ is the 3-D Cartesian coordinates of the pointing fingertip at t_n . In general, the predictive display demands reliable pointing finger tracking at a rate exceeding 30 Hz, as the majority of in-vehicle

pointing tasks can have durations in the range of $0.2 \text{ s} \leq T \leq 4 \text{ s}$ [8]. Figure 3 depicts three complete 3-D pointing trajectories, $Y_{1:T}$, collected in a car using an LM controller under three conditions, which visibly affect the pointing gesture.

Bayesian intent inference

Let $\mathbb{D} = \{\mathcal{D}_i : i = 1, 2, \dots, N\}$ be the set of N selectable items on the interactive display. While no assumptions are made about the layout of the icons in \mathbb{D} , each item is modeled as a distribution representing the extended regions in space of various shapes and sizes occupied by the corresponding GUI elements. For simplicity and computational efficiency, Gaussian distributions can be considered, and the i th item is modeled as $\mathcal{N}(\mu_i, C_i)$. The mean μ_i and covariance matrix C_i capture the 3-D location and the extent-orientation of \mathcal{D}_i , respectively. At the time instant t_k , the inference module in Figure 1 calculates the posterior probabilities for the N destinations,

$$\mathcal{P}(t_k) = \{p(\mathcal{D}_I = \mathcal{D}_i | Y_{1:k}) : i = 1, 2, \dots, N\}, \quad (2)$$

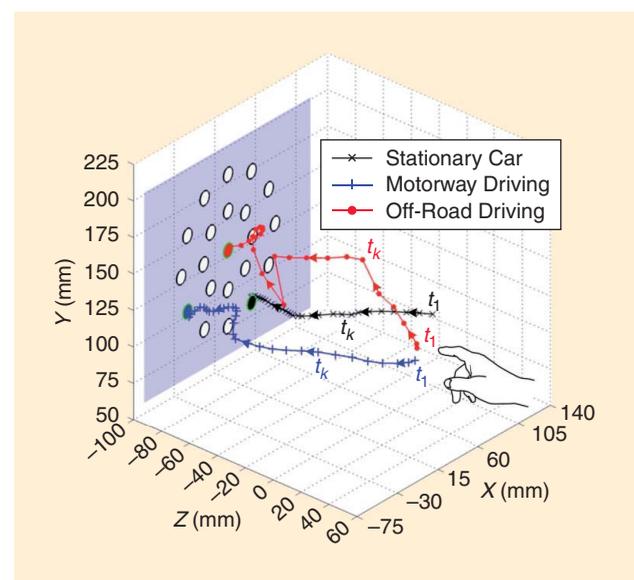


FIGURE 3. Recorded 3-D pointing fingertip tracks to select icons (circles) on an in-car touch screen (blue surface) under various driving/road conditions [10]; $t_1 < t_k$.

which represent the likelihood of each of the icons in \mathbb{D} being the unknown intended on-display end point $\mathcal{D}_I \in \mathbb{D}$. This uses the gesture tracker measurements $Y_{1:k}$ (and possibly other sensory data), i.e., the available partial pointing finger track at t_k whose extraction might require simple data sorting and associating routines. Each observation Y_k is assumed to be derived from an underlying pointing finger true (perturbation-free) latent state X_{t_k} , that can include its position, velocity, etc.

Within a Bayesian framework, we have

$$p(\mathcal{D}_I = \mathcal{D}_i | Y_{1:k}) \propto p(Y_{1:k} | \mathcal{D}_I = \mathcal{D}_i) p(\mathcal{D}_I = \mathcal{D}_i), \quad (3)$$

where the prior $p(\mathcal{D}_I = \mathcal{D}_i)$ on the selectable items (independent of $Y_{1:k}$ or the current pointing task) can be attained from relevant semantic and contextual information, such as selection frequency, GUI design, user profile, etc. This makes the adopted formulation particularly appealing as additional information, when available, can be easily incorporated. For example, the priors in (3) can be gradually and dynamically learned as the system is being used, starting from uninformative ones. Therefore, it is an adaptable probabilistic (belief-based) approach.

A prediction is performed at the arrival of each (or a few) new sensor observation(s). The inference module can use a number of low complexity, computationally efficient end-point predictors that are amenable to real-time implementation, given the limited computing resources and training data available in vehicles. The linear models discussed in the next section lead to a Kalman-filter-type implementation, combining end-point prediction and filtering out of noise induced by road/driving conditions. For severe perturbations, a separate statistical filter can be employed to remove the highly nonlinear gesture motion arising from perturbations.

Facilitation scheme and decision

To assist the selection task, the displayed interface may be modified at t_k , e.g., icons can be expanded/shrunk, colored/faded, or other [11], [12], [14], as per their probabilities $\mathcal{P}(t_k)$ in (2). Such facilitation strategies can require major modifications to legacy in-vehicle GUI designs and possibly the related software-hardware architectures. Their impact on the user experience in a split attention scenario (driving and interacting) is nontrivial and can be advised by experimental studies. For instance, unlike mouse pointing on a computer screen, constantly changing the in-car interface can increase visual demand to monitor the ongoing changes. A promising pointing facilitation scheme is midair selection, where the system autoselects the predicted intended on-screen item on behalf of the user, who does not need to physically touch the display surface. While midair selection can reduce the freehand pointing gesture duration and thus effort (visual, cognitive, and manual), its implementation entails only sending/reading

a select signal to/by the existing interface software module with minimal display overheads.

After inferring $p(\mathcal{D}_I = \mathcal{D}_i | Y_{1:k})$ at time t_k , the end point $\hat{\mathcal{D}}_I(t_k) \in \mathbb{D}$ of a freehand pointing gesture can be estimated (if needed) by minimizing the expected value of a cost function over all of the possible destinations in \mathbb{D} . This can be expressed by

$$\hat{\mathcal{D}}_I(t_k) = \arg \min_{\mathcal{D}^* \in \mathbb{D}} \sum_{i=1}^N C(\mathcal{D}^*, \mathcal{D}_i) p(\mathcal{D}_I = \mathcal{D}_i | Y_{1:k}), \quad (4)$$

where $C(\mathcal{D}^*, \mathcal{D}_i)$ is the cost of deciding \mathcal{D}^* as the destination, given that \mathcal{D}_I is the true intended on-display icon. If the binary decision criterion $C(\mathcal{D}^*, \mathcal{D}_i) = 1$ if $\mathcal{D}^* \neq \mathcal{D}_i$ and $C(\mathcal{D}^*, \mathcal{D}_i) = 0$ otherwise is used, it can be easily seen

that (4) leads to the maximum a posteriori (MAP) estimate; it implies that the most probable end point is deemed to be the intended on-display selectable icon. Within the Bayesian framework, more elaborate cost functions can be applied [23]; groups $\hat{\mathbb{D}}_q \subset \mathbb{D}$ rather than an individual icons may also be considered for expansion or fading purposes.

While the intuitive MAP estimate can be used to assess the suitability of the prediction model, it can produce fast fluctuating decisions during the pointing task.

This can be detrimental to midair selection due to the resultant false positives. In such cases, a simple decision rule can stipulate that the probability of an icon $p(\mathcal{D}_I = \mathcal{D}_i | Y_{1:k})$, specifically the one delivered by the MAP classifier, should exceed a certain threshold for a given duration of time before triggering an autoselection action.

Adaptable GUI and selection confirmation

The displayed interface implements seamlessly, in real time, the applied pointing facilitation scheme. If an on-screen item is selected or autoselected, the user can substantially benefit (i.e., in terms of reducing the visual workload) from a feedback confirming the selection action, e.g., an audible or haptic signal. For a predictive display with midair selection, the emerging ultrasonic midair haptic technology [24] presents itself as a suitable equivalent to the conventional on-screen haptic feedback, which is used in standard smartphone devices, with established benefits.

Additional sensory data

The availability of additional vehicle sensory data, such as suspensions travel data via the controller area network (CAN) bus or an on-board inertial measurement unit (IMU), can enable the intelligent predictive display system to establish the operating conditions, for instance, allowing it to determine whether the user input is perturbed or even estimate the level of noise present. It can then modify the applied statistical model by adapting its parameters or performing preprocessing

In the area of object tracking, knowing the destination of a tracked object not only leads to more accurate tracking results, but also offers vital information on intent, revealing potential conflict or threat.

prior to intent inference. Eye-gaze measurements can also offer valuable information on areas of interest on the display and can be used as an input modality in HCI, e.g., in [25]. Eye-gaze trackers are primarily utilized to examine the human performance behavior in a controlled setting, such as simulators, and a corpus of literature exists [26]. Obtaining accurate data from such a tracker, which is not head mounted, in a mobile vehicle can be challenging given the currently available commercial sensors. However, the fusion or simultaneous use of eye-gaze and pointing gesture data for an in-vehicle predictive display is a promising research area. In summary, if any additional information becomes available, it can be easily incorporated into the Bayesian framework via the priors $p(\mathcal{D}_I = \mathcal{D}_i)$, $\mathcal{D}_i \in \mathcal{D}$ or, alternatively, treated as a part of the measurements vector $Y_{1:k}$.

Bayesian end-point prediction

Given the available measurements $Y_{1:k}$ at t_k , determining the probability of each of the end points in \mathbb{D} being the intended destination requires calculating the observation likelihood $p(Y_{1:k} | \mathcal{D}_I = \mathcal{D}_i)$ conditioned on each end point, as stated in (3). The prior $p(\mathcal{D}_I = \mathcal{D}_i)$, which is independent of the current pointing task, is presumed to be available; here, for simplicity, all icons are assumed to be equally probable with $Pr(\mathcal{D}_I = \mathcal{D}_i) = 1/N$, $i = 1, 2, \dots, N$. The key problem in the intent prediction procedure is, therefore, that of evaluating the observation likelihood, i.e., the probability of having made a series of observations, under the assumption that the tracked object is ultimately heading to a given destination. This can be tackled by adopting an underlying motion model of the pointing finger, describing its trajectory on its journey toward the intended end point, and including an element of randomness in the followed track. This capitalizes on the premise that the motion of the pointing finger in 3-D is dictated by the intended icon on the display. Since the true destination \mathcal{D}_I is unknown a priori, N such models for each $\mathcal{D}_i \in \mathbb{D}$ are postulated, and the objective becomes calculating the likelihood of the observed partial pointing trajectory being drawn from a particular end-point-driven model. In other words, the destination that leads to a model that best explains $Y_{1:k}$ is assigned the highest probability of being \mathcal{D}_I , and vice versa.

According to the chain rule of probability,

$$p(Y_{1:k} | \mathcal{D}_I = \mathcal{D}_i) = p(Y_k | Y_{1:k-1}, \mathcal{D}_I = \mathcal{D}_i) \times p(Y_{1:k-1} | \mathcal{D}_I = \mathcal{D}_i), \quad (5)$$

where $p(Y_{1:k-1} | \mathcal{D}_I = \mathcal{D}_i)$ is the likelihood estimated at the previous time instant t_{k-1} . Thus, the observation likelihood in (5) can be calculated sequentially, i.e., with the arrival of each new sensor measurement of the pointing gesture, and determining the prediction error decomposition (PED),

$p(Y_k | Y_{1:k-1}, \mathcal{D}_I = \mathcal{D}_i)$, at t_k suffices. Next, we outline simple destination-driven models, including the BD approach introduced in [27] and [28]. We also show how sequential calculation of the PED can be performed, permitting the posterior probability distribution over intended end points in (2) to be calculated at each stage.

Modeling pointing movement

The pointing gesture movement toward an on-screen item is not deterministic. The person making the pointing gesture is capable of autonomous action and is in control of a complex motor system with numerous physical constraints, and he or she is likely to also be subjected to external motion, jolting, rolling, acceleration, and braking in a moving vehicle. Hence, models of the pointing finger movements, albeit driven by intent, are uncertain, and this can be captured by adopting stochastic models. This implies that the predictions of the tracked object motion are not single deterministic paths but are rather probabilistic processes, with the pointing finger

position at a future time expressed as a probability distribution in space. By adequately incorporating this uncertainty, relatively simple models of pointing finger motion can be used successfully to evaluate the corresponding observation likelihoods and the probabilities of $\mathcal{P}(t_k)$ in (2). It is emphasized here that the in-vehicle predictive display system objective is to infer the intent of the hand movement—it is not to accurately model the complex

human motor system. Thus, an approximate motion model that enables reliably determining the destination of a freehand pointing gesture is sufficient.

Calculating the transition density of a stochastic model, e.g., between two successive observation times t_{k-1} and t_k , is required to condition the tracked pointing finger state X_t (e.g., position, velocity, etc.) on a nominal end point \mathcal{D}_i . Continuous-time motion models are a natural choice, where the tracked object's dynamics are represented by a continuous-time stochastic differential equation (SDE). This SDE can be integrated to obtain a transition density over any time interval. Although numerous models for object tracking exist, the class of Gaussian linear time invariant (LTI) models for the evolution of X_t is utilized by the in-vehicle predictive display, as they lead to a low-complexity inference procedure (unlike non-linear and/or non-Gaussian models). This class includes many models used widely in tracking applications, for example, the (near) constant velocity (CV) and linear destination reverting (LDR) models highlighted next, in addition to other Gaussian LTI models that can describe higher-order kinematics (acceleration, jerk, etc.) [22].

While the system governing the pointing finger dynamics is assumed not to change over time, it does depend here on the intended end point $\mathcal{D}_i \in \mathbb{D}$, which intrinsically drives the pointing motion. Conditioned on knowing this end point, e.g., the i th GUI icon \mathcal{D}_i , and integrating the Gaussian LTI model,

Destination prediction can be viewed as a means to assist planning and decision making at a system level higher than that of established conventional sensor-level tracking algorithms.

the relationship between the system state at times t and $t + h$ can be written as

$$X_{i,t+h} = F(h, \mathcal{D}_i)X_{i,t} + M(h, \mathcal{D}_i) + \varepsilon_t, \quad (6)$$

where $\varepsilon_t \sim \mathcal{N}(0, Q(h, \mathcal{D}_i))$ is the dynamic noise embodying the randomness in the motion model. The matrices F and Q as well as the vector M , which together define the state transition from one time to another, are functions of the time step h and, notably, the destination $\mathcal{D}_i \in \mathbb{D}$. Thereby, N such models are constructed to establish the end point of the pointing gesture.

The k th observation, for example, the pointing finger position as provided by the gesture tracking device, is also modeled as a linear function of the time t_k state perturbed by additive Gaussian noise,

$$Y_k = GX_{i,t_k} + \nu_k, \quad (7)$$

where G is a matrix mapping from the hidden state to the observed measurement and $\nu_n \sim \mathcal{N}(0, V_n)$. For instance, if the gesture tracker provides the pointing finger positions directly and the system state includes only position, then G is a 3×3 identity matrix. The noise covariance can be utilized to set the level of noise in each of the x , y , and z axes as per the gesture tracker specifications; e.g., a time-of-flight-based tracker such as the SoftKinetic DepthSense camera exhibits higher inaccuracies in observations along the depth axis. It is noted that no assumption is made about the observation arrival times t_k and irregularly spaced, asynchronous measurements can naturally be addressed within this formulation. The system structure, for each nominal end point \mathcal{D}_i , is depicted graphically in Figure 4, where the destination \mathcal{D}_i influences the end-point-driven state at all times.

Among linear Gaussian models, linear destination reverting models, such as the mean reverting diffusion (MRD) and equilibrium reverting velocity (ERV) models, make particularly suitable candidates for the pointing finger motion in (6), as discussed in [10]. Their state evolution explicitly incorporates the destination information. For example, the governing SDE for the MRD model is given by $dX_{i,t} = \Lambda(d_i - X_{i,t})dt + \sigma dw_t$. It indicates an attraction of the motion toward the location of

destination d_i (e.g., the mean of the Gaussian distribution representing \mathcal{D}_i), with Λ (a design parameter) capturing the strength of this reversion for each axis in 3-D and w_t being a Wiener process. While the MRD is based on a multivariate Ornstein–Uhlenbeck process [29] and the system state includes only the position information in 3-D, the state of the ERV model proposed in [10] additionally includes the velocity of the pointing finger, in 3-D, driven by the end point. This facilitates modeling pointing velocity profiles like those shown in Figure 2(b). Integrating the SDE of the MRD and ERV results in (6), each with specific F , M , and Q matrices.

During a pointing task, the path of the pointing finger, albeit random, must end at the intended destination at time T (i.e., the pointing finger reaches its end point on the display). This can be modeled by an artificial prior probability distribution for X_T corresponding to the geometry of the destination; alternatively, it can be treated as a pseudo-observation at T . To maintain the linear Gaussian structure of the system in (6) and (7), this distribution is assumed to be Gaussian, such that $p(X_T | \mathcal{D}_i = \mathcal{D}_i) = \mathcal{N}(X_T; a_i, \Sigma_i)$; see [28] for a discussion on this construct. The mean vector a_i specifies the constrained system state at the destination, whereas Σ_i is a covariance matrix of the appropriate dimension. For instance, for the MRD model, in which only pointing finger position is considered, $a_i = \mu_i = d_i$ representing the location-center of the destination in 3-D. In the case of the ERV model, defining the final state distribution also involves specifying a distribution of the pointing finger velocity at the end point. A large-scale prior covariance can be used to model the uncertainty in this; however, certain properties might be assumed, e.g., relatively high velocity in the direction toward the screen.

Exploiting the artificial prior on the distribution of X_T requires that the state of the motion models in (6) to be conditioned not only on $\mathcal{D}_i \in \mathbb{D}$ but also on the arrival time T . Including this permits the posterior of the system state at time t_k to be expressed as $p(X_{t_k} | Y_{1:k}, T, \mathcal{D}_i = \mathcal{D}_i)$, and the sought observation likelihood in (5) is subsequently given by $p(Y_{1:k} | T, \mathcal{D}_i = \mathcal{D}_i)$ after k measurements. The inclusion of the prior on X_T in the motion model changes the system dynamics (even for MRD and ERV models), where the predictive distribution of the pointing finger state changes from a fully random walk to a bridging distribution (BD), terminating at the end point. This encapsulates the long term dependencies in the pointing finger trajectory due to premeditated actions guided by intent. Since the intended destination is not known, N such bridges are constructed, one per nominal end point. Consequently, all Gaussian linear models, including the nondestination reverting ones, whose dynamic models are not dependent on \mathcal{D}_i like Brownian motion (BM) and CV, can be utilized for destination prediction within the presented Bayesian framework. This technique of conditioning on the end point is dubbed *BD-based inference*.

Intent inference: Sequential likelihood evaluation

We recall that the primary objective of the intent inference routine is to determine the observation likelihoods $p(Y_{1:k} | \mathcal{D}_i = \mathcal{D}_i)$, $\mathcal{D}_i \in \mathbb{D}$, at t_k , rather than the posterior

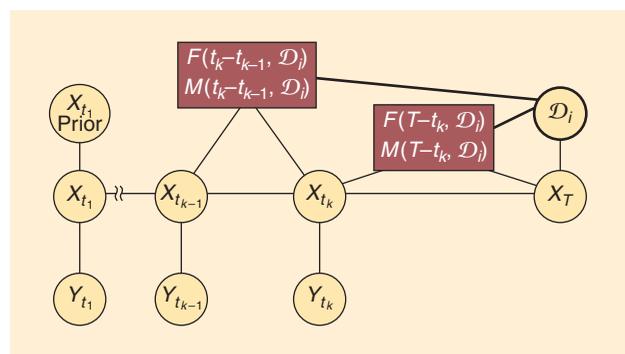


FIGURE 4. The system graphical structure; end point \mathcal{D}_i acts as a prior and affects the state transition.

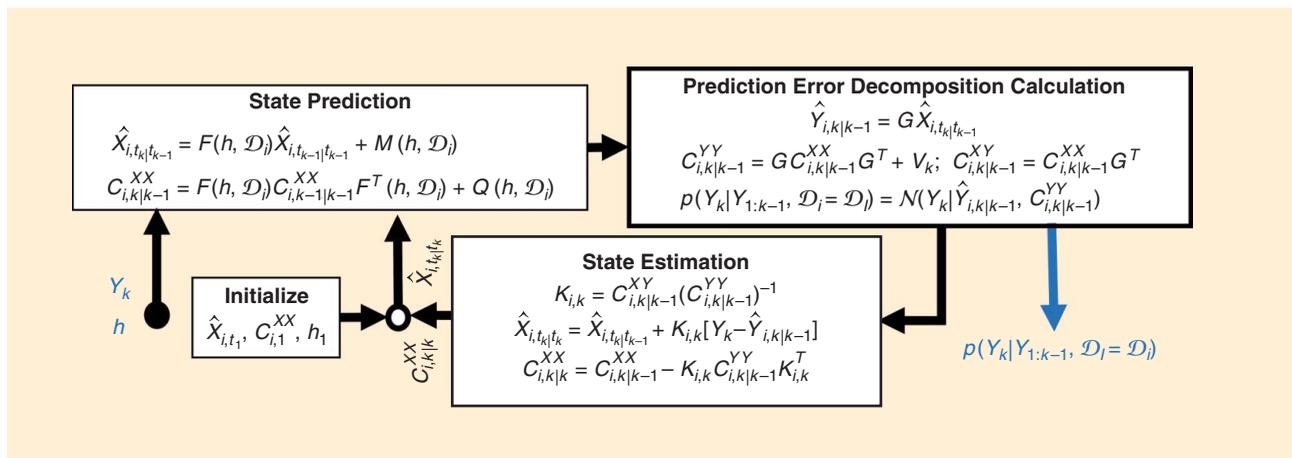


FIGURE 5. The Kalman filter for sequentially evaluating the PED for end point \mathcal{D}_i at the arrival of observation Y_k ; state prediction at t_k relies on the state estimation results, including covariance $C_{i,k-1|k-1}^{XX}$, from the previous time step and $h = t_k - t_{k-1}$.

distribution of the system state X_{t_k} , as in traditional tracking applications [22]. Nonetheless, the latent state estimation, which might be relevant in certain scenarios, is addressed below. Based on (6) and (7), a classical Kalman filter can be employed to sequentially calculate the prediction error decomposition in (5) as depicted in Figure 5 and, thereby the sought observation likelihood for the current set of measurements $Y_{1:k}$ conditioned on \mathcal{D}_i . The computationally efficient Kalman filter is particularly desirable since running, concurrently, multiple Kalman filters for all $\mathcal{D}_i \in \mathbb{D}$ is plausible in real-time, even in settings where limited computing power is available. This solution is also amenable to parallelization.

For the bridging approach, it is shown in [27] and [28] how the PED and observation likelihood in (5) from each constructed bridge, i.e., conditioned on T and \mathcal{D}_i , can be estimated using a modified Kalman filter. As the true arrival time T is unknown a priori in practice, approximating

$$p(Y_{1:k}|\mathcal{D}_i = \mathcal{D}_i) = \int_{T \in \mathcal{T}} p(Y_{1:k}|T, \mathcal{D}_i = \mathcal{D}_i)p(T|\mathcal{D}_i = \mathcal{D}_i)dT, \quad (8)$$

is necessary, where $p(T|\mathcal{D})$ is the prior distribution of arrival times at destination \mathcal{D}_i and \mathcal{T} is the time interval of possible arrival times T . In the simplest case, arrivals might be assumed at some specific future time. This is a crude approximation; nevertheless, is often quite effective [28]. To improve inference accuracy (and possibly also to learn about expected arrival time), arrivals can be modeled as having a prior distribution, such as being expected uniformly within some time period $[t_a, t_b]$, giving $p(T|\mathcal{D}) = \mathcal{U}(t_a, t_b)$. In this case, numerical quadrature, for example, via Simpson’s rule, can be applied. Although BD-based intent inference involves running multiple Kalman filters, and, hence, is more computationally demanding, it can significantly improve the end-point inference capability of a predictive display and leads to a more robust performance.

In summary, the introduced modeling approach for inferring as early as possible the item that the user intends to select on the display using the freehand pointing gesture is generic. Most importantly, it offers considerable flexibility in terms of catering to various sensing technology specifications (e.g., observation error) as well as adaptability in terms of adjusting the motion model parameters. The approach is simple and relatively computationally efficient, which makes it suitable for the requirements of an automotive environment. In the developed predictive display prototype (an optimized C# implementation of the system in Figure 1 on a typical automotive computing platform), prediction with Kalman filtering was tested with up to $N = 64$ destinations and an observations data rate ≥ 30 Hz without any noticeable delays in the system response in terms of the pointing facilitation routine.

Handling perturbed pointing trajectories

When the user input is perturbed in a moving vehicle due to the road and driving conditions, the predictive display system can handle noisy freehand pointing gestures by setting the noise covariance in the motion model in (6) relative to the measured (experienced) in-vehicle vibrations/accelerations. This conforms with the modeling assumptions, and a higher covariance corresponds to having less certainty in the inferred end-point-driven latent state $X_{i,t}$, i.e., pointing finger position, velocity, etc. This technique is suitable for low to medium perturbation levels that can be represented by Gaussian noise, for instance, driving on smooth to moderately bumpy paved roads. The output of the filters, calculating the posterior of each nominal destination $p(\mathcal{D}_i = \mathcal{D}_i | Y_{1:k})$ at t_k , can be used to estimate the posterior probability of the system latent state X_{t_k} , including the perturbation-free pointing finger position. This is given by the Gaussian mixture

$$p(X_{t_k} | Y_{1:k}) = \sum_{i=1}^N p(X_{i,t_k} | Y_{1:k})p(\mathcal{D}_i = \mathcal{D}_i | Y_{1:k}), \quad (9)$$

where $p(X_{i,t_k} | Y_{1:k})$ pertains to the i th destination and is also calculated by the Kalman filter.

The assumption of Gaussian noise in a motion model can be overly restrictive in highly perturbed environments, e.g., driving on a rough terrain or a badly maintained road, since the pointing hand/finger can move in a highly erratic manner. It can exhibit sudden unintentional noise-related movements or jolts, as can be seen in Figure 3 for off-road driving. In such scenarios, the perturbations present can be treated as an additional nonlinear random jump process, denoted by P_t in the motion model, causing sudden large changes in the pointing finger position and velocity. For example, this can be modeled by the mean-reverting jump-diffusion velocity process whose SDE is given by

$$d\dot{P}_t = \sigma_p dW_{2,t} + \sigma_J dJ_t - \lambda_1 \dot{P}_t dt, \quad (10)$$

such that dJ_t is the instantaneous change in the jump process $J_t = \sum_{i=1}^{\tau_t} \rho_i$, with $\rho_i \sim \mathcal{N}(0, 1)$, τ_t is the number of jumps in $[0, t]$ governed by a Poisson distribution, and the next jump time τ is set by an exponential distribution [18]. Likelihood estimation for such motion models relies on sequential Monte Carlo (SMC), particle, filtering [30], which is computationally costly and approximate compared to the original models in (6) using Kalman filtering, even in the efficient Rao–Blackwellized form [22], [30]. A practical alternative to applying this expensive inference procedure N times, one per destination, is to apply the SMC filtering once as a preprocessing stage prior to the destination prediction routine. The preprocessing objective is to remove the most severe effects of large jolts from the gesture tracker observations $Y_{1:k}$ at t_k and allow the utilization of the original linear motion models for intent inference [10], [28]. This approach represents a compromise between the better filtering results of the jump model in a high-perturbations environment, and the computational efficiency of the original models.

Applying a preprocessing SMC filter or dynamically adjusting the motion model covariance can be guided by additional sensory data, such as changes in the suspension height (by probing the vehicle CAN bus), IMU accelerometer, front-facing cameras, etc. These can reliably measure the level of accelerations and vibrations experienced in the vehicle. Additionally, the filtered freehand pointing gesture can be used not only for pointing, but also for general gesture-based interactions.

Performance analysis: Empirical results

The performance of the intelligent predictive display concept is assessed here using data collected in two cars (a Jaguar XK and a Range Rover) instrumented with the system in Figure 1 under various road and driving conditions, specifically when the vehicle is 1) stationary, 2) driven over a well-maintained road (i.e., motorway) at varying speeds (30–70 mi/h), and 3) driven on a badly maintained road

The assumption of Gaussian noise in a motion model can be overly restrictive in highly perturbed environments, e.g., driving on a rough terrain or a badly maintained road, since the pointing hand/finger can move in a highly erratic manner.

with rutted and potholed surfaces with random patches and raised/sunken man-hole covers, where mild to severe in-car perturbations are experienced. A Leap Motion sensor is used to track, in real time, the freehand pointing gestures (pointing fingertips), and an experimental GUI is displayed on an 11.5-in touch screen mounted to the car's dashboard. Two videos available as supplementary material with this article in IEEE *Xplore* show an early prototype of a predictive touch-screen system; see [http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?](http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=79)

[punumber=79](http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=79). The interface has $N = 21$ selectable circular icons, each of width $W \leq 2$ cm that are approximately 2 cm apart in a circular formation, identical to that in Figure 3; the detailed setup is described in [8]. Similar to the Fitts' law task in HCI, one randomly chosen GUI item is highlighted at a time, and the user is expected to select it via a freehand pointing gesture. To maintain an objective testing procedure, all possible end points in \mathbb{D} are assumed to be equally probable, $\Pr(\mathcal{D}_t = \mathcal{D}_i) = 1/N, i = 1, 2, \dots, N$. Maximization of the likelihood function $\prod_{j=1}^J p(Y_{1:n}^j | \mathcal{D}_t = \mathcal{D}_i, \Omega)$ for a sample of J typical full pointing finger trajectories is used to set the motion model parameters Ω , and thus constitutes training for the system. Next, the performance results of several Bayesian predictors and an in-car prototype system are examined. It is emphasized that predictors have no knowledge of the user intent in any of the experiments described next.

End-point prediction performance

To examine the prediction accuracy throughout the pointing task, from its start at t_1 until touching the display surface at time T , 50 a priori recorded in-car full pointing gestures are used; no pointing facilitation routine is applied. The inference performance is evaluated in terms of 1) the ability to determine the intended on-screen icon via a MAP estimate $\hat{\mathcal{D}}(t_k) = \operatorname{argmax}_{\mathcal{D}_i \in \mathbb{D}} p(\mathcal{D}_t = \mathcal{D}_i | Y_{1:k})$, i.e., how early the predictor assigns the highest probability to true end point \mathcal{D}^+ and 2) the aggregate inference success, i.e., the proportion of the total pointing gesture (in time) for which the predictor correctly inferred \mathcal{D}^+ . The success is defined by $S(t_n) = 1$ if $\hat{\mathcal{D}}(t_n) = \mathcal{D}^+$ and $S(t_n) = 0$ otherwise, for observations at times $t_n \in \{t_1, t_2, \dots, T\}$. While $J = 5$ pointing trajectories are used for training, the prior on the distribution of the durations of typical in-car pointing tasks, $p(T | \mathcal{D}_t = \mathcal{D}_i)$, for the bridging-distribution predictors is obtained from the experimental study in [8]. It is noted that utilizing 10% of the available tracks to set the model parameters is aimed at demonstrating the low training requirement of the applied state-space-modeling-based inference approach. This feature is highly desired in an automotive context as discussed in the “Background” section. However, as the driver/passenger uses the predictive display, the system can

refine the applied model parameters from the larger available data set(s). This can result in a more accurate modeling and prediction procedure.

In Figure 6, the linear destination reverting, Brownian motion, and constant velocity models with the bridging prior, notated by MRD-BD, ERV-BD, BM-BD, and CV-BD, are assessed. A mean reverting diffusion model without bridging, MRD, is also examined. Figure 6 also depicts the outcome of the probabilistic nearest neighbor (NN), which assigns the highest probability to a GUI item closest to the current position of the pointing fingertip as per $p(Y_n | \mathcal{D}_I = \mathcal{D}_i) = \mathcal{N}(Y_n; d_i, C_{NN})$, and bearing angle (BA) where $p(Y_n | Y_{n-1}, \mathcal{D}_I = \mathcal{D}_i) = \mathcal{N}(\theta_{i,n}; 0, \sigma_{BA}^2)$ [10]. The latter assumes a minimal cumulative angle to the destination located in 3-D at d_i ; C_{NN} is the covariance of the multivariate Gaussian distribution and $\theta_{i,n} \triangleq \angle(Y_n - Y_{n-1}, d_i)$ is the angle to $\mathcal{D}_i \in \mathcal{D}$.

Figure 6 illustrates that the BD-based inference models, CV-BD and ERV-BD, achieve the earliest successful predictions, since they capture the importance of the velocity component. This is particularly visible in the first 70% of the pointing task in Figure 6(a), where a pointing facilitation scheme can be most effective. Destination prediction toward the end of the pointing gesture can have limited impact, since by that stage the user would have already dedicated the necessary attention/effort to execute the selection task. The performance of all depicted predictors generally improves as the pointing finger is closer to the display. This is particularly visible for the NN model, which is built on the premise that the pointing finger is closest to the intended end point. An exception is the BA model, since the reliability of θ_n as an intent measure declines as $t_n \rightarrow T$. Overall, this figure shows that probabilistic predictors can successfully infer the intended destination on the display remarkably early in the freehand pointing gesture. For example, in 60% of cases, the bridged ERV model, ERV-BD, can infer the true intent only 40% into the pointing gesture (with overall correct decision exceeding 65%) thus, it can reduce pointing time-effort by over 60%.

The gains of combining the MRD motion model with the bridging method are noticeable in Figure 6(a). This is due to the ability of bridging technique (the prior on X_T) to reduce the sensitivity of LDR models to variability in the processed tracks; it tapers the system sensitivity to parameter estimates and the parameter training requirements.

Real-time results from a prototype system

Here, results from a pilot user study with 20 participants are presented. While none of the participants has used an intent-aware display before, the study employs a prototype in-car intelligent predictive touch-screen system that performs intent inference in real time and seamlessly implements the midair selection facilitation scheme as discussed in the section “Facilitation Scheme and Decision”; see the supplementary material in IEEE *Xplore* for a demonstration. An audible cue, i.e., a short ping sound signal, is produced by the predictive display to con-

The reviewed concept of intelligent predictive displays in this article presents itself as a promising smart HMI technology.

firm to the user that an interface icon has been autoselected. The subjective workload of interacting with an in-vehicle touch screen with and without the predictive functionality is recorded using the NASA TLX test [31], which is widely utilized in HMI-HCI studies. It requires the participant to complete a questionnaire to rate and weight the mental,

physical, and temporal demand as well as performance, effort, and frustration experienced when carrying out the in-vehicle pointing tasks. The durations T of accomplishing selection tasks in the trials are also assessed. This can be viewed as an objective measure of the effort involved.

When the predictor is off, the trial is a classical experiment of interacting with a conventional touch screen, where the user has to touch the display surface to select a GUI icon. With the prediction and midair selection functionality on, the intent-aware predictive touch screen often executes the selection action for the user. An autoselection action is triggered at time $t_k \leq T$ once the calculated probability for a given GUI icon, as per the estimated $p(\mathcal{D}_I = \mathcal{D}_i | Y_{1:k})$, exceeds a set threshold for a predefined period of time (on average, threshold $\gamma = 0.55$

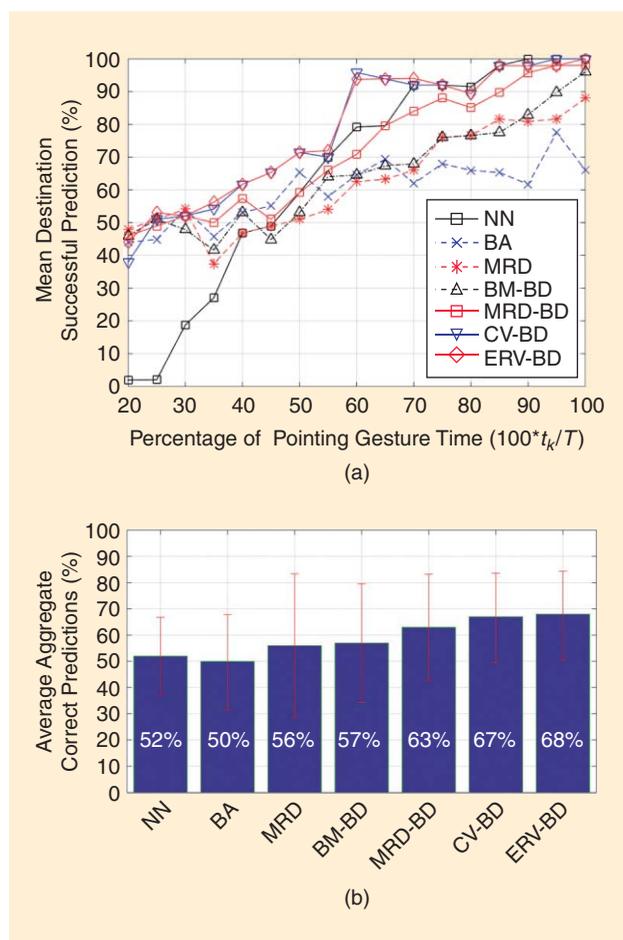


FIGURE 6. The end-point inference performance with a MAP estimate as a function of the percentage of pointing time [28]. (a) The mean successful destination inference. (b) The gesture portion (in time) with successful predictions.

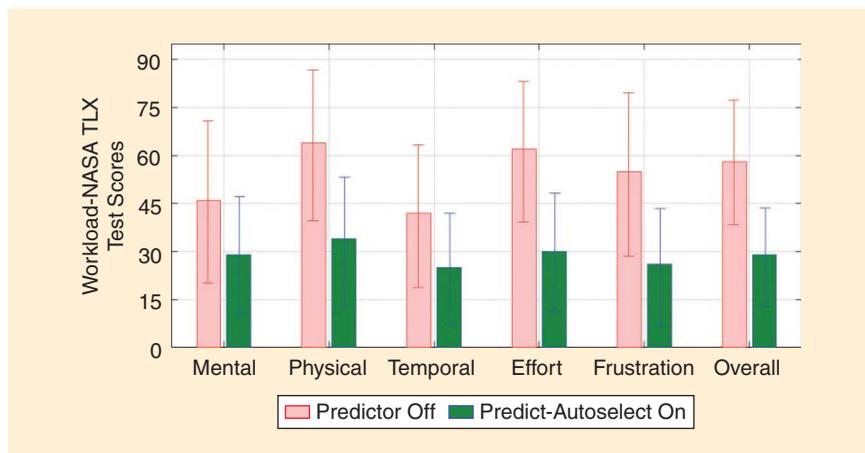


FIGURE 7. The mean workload of using an in-car display with and without the predictive functionality for 20 participants.

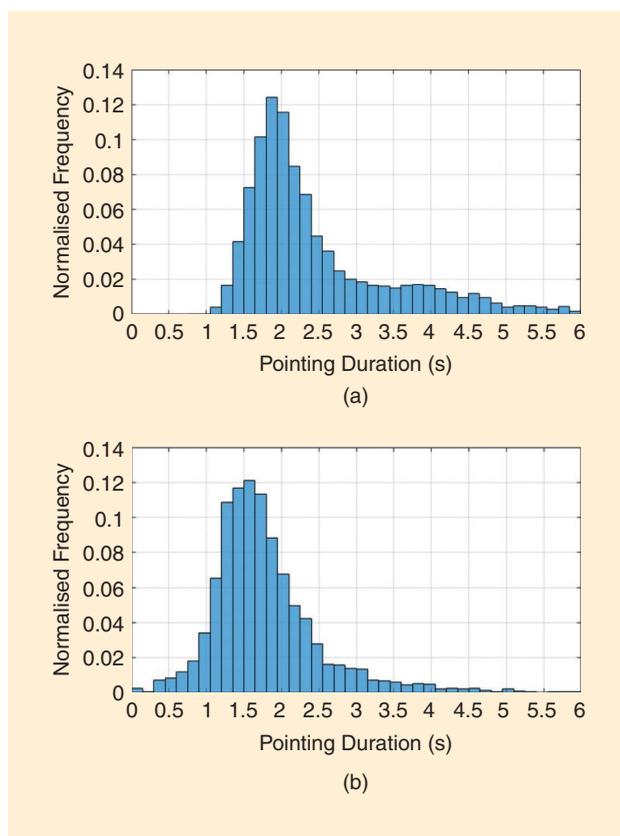


FIGURE 8. The pointing time (in seconds) with and without prediction-autoselection functionality for 20 participants. (a) Predict-autoselect off; $\mu = 2.63$ s and $\sigma = 1.34$. (b) Predict-autoselect on; $\mu = 1.82$ s and $\sigma = 0.84$.

and its duration $T_S = 65$ ms are set empirically). When this prediction certainty requirement is not met or the pointing finger is not detected, the user can continue pointing until he/she touches (and selects) the intended interface icon. Since the system is not aware of the user intent, any erroneous selection of the unintended GUI icon will lead to a longer pointing time

and higher subjective workload, e.g., a higher frustration score.

Figure 7 shows that the interactions subjective overall workload declines by over 47% when employing the predictive display system in Figure 1 with midair selection, which is a substantial reduction. Figure 8 depicts the normalized histogram of pointing tasks duration T for more than 8,000 selection tasks for all 20 participants. This figure illustrates that T is reduced when the prediction-autoselection functionality is on. In particular, the histogram in Figure 8(b) is visibly shifted to the left with smaller durations being more frequent and high values (indicating lengthy effortful

pointing gestures) less recurring. On average, the introduced predictive solution reduces the duration of accomplishing an on-screen selection task via a freehand pointing gesture by approximately 30.75%. Higher reductions in the pointing time can be achieved (see Figure 6) by relaxing the requested prediction certainty (threshold or its duration) at the expense of, possibly, increasing the number of false autoselections. This can have a negative impact on the user experience and system acceptance. It is a tradeoff that has to be taken into account, and the decision criterion can be adaptively changed based on the user requirements and the controlled IVIS functionality or the displayed GUI.

Remarks on results

Since interactions with displays in modern vehicles are prevalent [2], small improvements in the pointing task efficiency, even reducing its duration by a few milliseconds, can have significant aggregate benefits on the user experience, notably for drivers. Therefore, the overviewed predictive solution can substantially reduce the effort and distraction of using in-vehicle interactive displays. However, further experimental evaluation is required for other pointing facilitation schemes, in lieu of midair selection that involves taking an action on behalf of the user. Additionally, devising a principled approach to setting the decision criterion for autoselection according to the general cost minimization problem in (4) is an open research question.

Conclusions

Recent advances in sensing, data storage, and communications technologies have led to the introduction of new smart vehicle functionalities and services aimed at offering personalized, more pleasant, and safer driving experience. Nevertheless, little attention is often paid to the human-machine interface aspect of these functionalities, for instance, interacting, controlling, and customizing them. Such interactions can be highly effortful and distracting, especially for drivers,

with potential safety consequences. The reviewed concept of intelligent predictive displays in this article presents itself as a promising smart HMI technology. It can significantly reduce the effort and distractions associated with using an in-vehicle interactive display, which typically serves as a gateway to the available IVIS and services. This solution, whose cornerstone is suitable statistical signal processing algorithms, can also allow interaction with displays that do not have a physical surface, such as HUDs for augmented reality and projections of 3-D interfaces; such displays are poised to proliferate rapidly in the automotive environment in the near future.

Within the introduced general Bayesian framework, additional sensory or semantic data, if available, can be easily incorporated to enhance the prediction capabilities of the intelligent display and its handling of perturbed freehand pointing gestures due to road and driving conditions. The perturbations filtering aspect of this solution can be beneficial to general gesture-recognition-based interfaces in vehicles, not only pointing. Moreover, the predictive system can offer additional flexibilities in terms of the interface design and display placement in the vehicle interior, as users might only need to reach (not necessarily touch) the display, with the midair selection scheme. This can be viewed in the context of inclusive design and ergonomics, where the display response or operation mode can be tailored to the user profile and motor abilities. Predicting the intended end point of a freehand gesture can extend, beyond the touch screen to other items within the vehicle, such as the various mechanical controls.

Although a number of predictors that are based on Gaussian motion models were discussed here, several other probabilistic approaches can be employed within the presented Bayesian formulation, such as interacting multiple models [22], stochastic context-free grammars [16], and other destination-aware tracking algorithms. While the presented empirical results testify to the efficacy of the intelligent predictive display system, this solution can benefit from future advancements in in-vehicle sensing technology, probabilistic intent inference algorithms, Bayesian decision strategies, fusion of multiple sensory data (not only gesture), and others. This article serves as an impetus for further research into using signal processing or machine learning techniques to alleviate the effort and attention required to interact with smart infotainment, connectivity, and safety services in vehicles.

Supplemental material

This article has supplementary downloadable material available at <http://ieeexplore.ieee.org>. The material includes two videos demonstrating the predictive display concept. Contact bia23@cam.ac.uk with any questions pertaining to this work.

Predicting the intended end point of a freehand gesture can extend, beyond the touch screen, to various other items within the vehicle, such as the various mechanical controls.

Video titled "Predictive Pointing: Prediction Results in Real Time"

A video demonstrating the prediction results calculated, in real time, by an in-vehicle intent-aware display for a few typical in-car freehand pointing gestures. The system has no prior knowledge of the intended on-screen item and no pointing facilitation scheme is applied (i.e., no midair selection).

Video titled "Predictive Touchscreen with Midair Selection"

A video that shows an early prototype of the predictive display system, with midair selection, operating in a laboratory set-up and in a moving vehicle.

Acknowledgment

We would like to thank Jaguar Land Rover for funding this work under the Centre for Advanced Photonics and Electronics agreement and supporting the data collection on their test-track at the JLR Gaydon Centre, United Kingdom.

Authors

Bashar I. Ahmad (bia23@cam.ac.uk) received his B.Eng. (Hons.) and Ph.D. degrees in electronic engineering, in 2007 and 2011, respectively, from the University of Westminster, United Kingdom. Currently, he is a senior research associate in the Signal Processing and Communications Laboratory, Engineering Department, Cambridge University, United Kingdom, and a fellow of Wolfson College, Cambridge. Prior to joining Cambridge University, he was a post-doctoral researcher in the Signal Processing and Communications group at Imperial College London. His research interests include statistical signal processing, multimodal human computer interactions, sampling theory, and cognitive radio.

James K. Murphy (jm362@cam.ac.uk) received his M.A. degree in computer science in 2003 and his Ph.D. degree from Cambridge University, United Kingdom, in 2013. He received his M.S. degree in mathematical modeling from Oxford University, United Kingdom, in 2004. Currently, he is a research associate in the Signal Processing and Communications Laboratory, Engineering Department, Cambridge University. Previously, he worked as a senior analyst with a quantitative finance consulting firm. His research interests include Monte Carlo methods for inference in univariate and multivariate time series problems.

Simon J. Godsill (sjg30@cam.ac.uk) received his B.A. and Ph.D. degrees from the University of Cambridge, United Kingdom, in 1989 and 1994, respectively. Currently, he is a professor of statistical signal processing in the Engineering Department, Cambridge University. His research interests include Bayesian and statistical methods for signal processing, Monte Carlo algorithms for Bayesian problems, modeling and enhancement of audio and musical signals, tracking, and high-frequency financial data. He has

published extensively in journals, books, and conferences. He was an associate editor of *IEEE Transactions on Signal Processing* and *Bayesian Analysis*, and he is a member of the IEEE Signal Processing Theory and Methods Committee. In 2002, he coedited a special issue of *IEEE Transactions on Signal Processing* on Monte Carlo methods in signal processing, and he has organized many conference sessions on related themes.

Patrick M. Langdon (pml24@cam.ac.uk) received his B.S. degree (Hons.) in psychology in 1984 and his Ph.D. degree in experimental psychology in 1989 from the University of Sheffield, United Kingdom. He is a principal research associate in the Engineering Department, Cambridge University, United Kingdom, and a lead researcher in inclusive design within the Engineering Design Centre. He has originated numerous collaborative research projects in design for inclusion and human-machine interaction (HMI) since joining the department in 1997. He has published extensively in various formats, as author, issue editor, and coauthor. Currently, he is the principal investigator or coinvestigator of three industrial projects, one commercial collaboration in automotive, and a coinvestigator of a four-year Engineering and Physical Sciences Research Council project on HMI for autonomous, smart, and connected control.

Robert W. Hardy (rhardy@jaguarlandrover.com) received his undergraduate degree (2006) in computer science, and his M.S. (2007) and Ph.D. (2010) degrees in the field of mobile and ubiquitous systems, all from Lancaster University, United Kingdom. He joined Jaguar Land Rover, Coventry, United Kingdom, in 2011 and is a lead research engineer within the Human-Machine Interface research team. He worked as a research associate at Lancaster University on a three-year industrial project with the Japanese phone operator, NTT Docomo, working in the area of near-field communications. His research focus at Jaguar Land Rover primarily covers interface design, head-up display, vision systems, and haptics.

References

- [1] R. Bishop, *Intelligent Vehicle Technology and Trends*. Norwood, MA: Artech House, 2005.
- [2] C. Harvey and N. A. Stanton, *Usability Evaluation for In-vehicle Systems*. Boca Raton, FL: CRC, 2013.
- [3] Tesla Motors. (2016, 18 Mar.) An evolution in automobile engineering (ModelS) [Online]. Available: www.teslamotors.com/models
- [4] K. Bark, C. Tran, K. Fujimura, and V. Ng-Thow-Hing, "Personal Navi: Benefits of an augmented reality navigational aid using a see-thru 3D volumetric HUD," in *Proc. Int. Conf. Automotive User Interfaces and Interactive Vehicular Applications*, 2014, pp. 1–8.
- [5] N. Broy, M. Guo, S. Schneegass, B. Pflieger, and F. Alt, "Introducing novel technologies in the car—conducting a real-world study to test 3D dashboards," in *Proc. 7th Int. Conf. Automotive User Interfaces and Interactive Vehicular Applications*, 2015, pp. 179–186.
- [6] (2016, 18 Mar.) Jaguar Land Rover. XE in-car technology [Online]. Available: <http://www.jaguar.co.uk/jaguar-range/xe/features>
- [7] N. Goode, M. G. Lenné, and P. Salmon, "The impact of on-road motion on BMS touch screen device operation," *Ergonomics*, vol. 55, no. 9, pp. 986–996, 2012.
- [8] B. I. Ahmad, P. M. Langdon, S. J. Godsill, R. Hardy, L. Skrypchuk, and R. Donkor, "Touchscreen usability and input performance in vehicles under different road conditions: An evaluative study," in *Proc. Int. Conf. Automotive User Interfaces and Interactive Vehicle Applications*, 2015, pp. 47–54.
- [9] S. G. Klauer, T. A. Dingus, V. L. Neale, J. D. Sudweeks, and D. J. Ramsey, "The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data," National Highway Traffic Safety Administration, Washington, D.C., DOT HS 810 5942006, 2006.
- [10] B. I. Ahmad, J. K. Murphy, P. Langdon, S. Godsill, R. Hardy, and L. Skrypchuk, "Intent inference for pointing gesture based interactions in vehicles," *IEEE Trans. Cybern.*, vol. 46, no. 4, pp. 878–889, 2016.
- [11] E. Lank, Y.C. N. Cheng, and J. Ruiz, "Endpoint prediction using motion kinematics," in *Proc. SIGCHI Conf. Human Factors Computing Systems*, 2007, pp. 637–646.
- [12] M. J. McGuffin and R. Balakrishnan, "Fitts' law and expanding targets: Experimental studies and designs for user interfaces," *ACM Trans. Comput. Hum. Interact.*, vol. 12, no. 4, pp. 388–422, 2005.
- [13] B. Ziebart, A. Dey, and J. A. Bagnell, "Probabilistic pointing target prediction via inverse optimal control," in *Proc. 2012 ACM Int. Conf. Intelligent User Interfaces*, 2012, pp. 1–10.
- [14] P. T. Pasqual and J. O. Wobbrock, "Mouse pointing endpoint prediction using kinematic template matching," in *Proc. SIGCHI Conf. Human Factors Computing Systems*, 2014, pp. 743–752.
- [15] E. Ohn-Bar and M. M. Trivedi, "Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations," *IEEE Trans. Intell. Transport. Syst.*, vol. 15, no. 6, pp. 2368–2377, 2014.
- [16] M. Fanaswala and V. Krishnamurthy, "Detection of anomalous trajectory patterns in target tracking via stochastic context-free grammars and reciprocal process models," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 1, pp. 76–90, 2013.
- [17] M. Goel, L. Findlater, and J. Wobbrock, "Walktype: Using accelerometer data to accommodate situational impairments in mobile touch screen text entry," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, 2012, pp. 2687–2696.
- [18] B. I. Ahmad, J. K. Murphy, P. M. Langdon, and S. J. Godsill, "Filtering perturbed in-vehicle pointing gesture trajectories: Improving the reliability of intent inference," in *Proc. IEEE Int. Workshop Machine Learning for Signal Processing*, 2014, Reims, France, pp. 1–6.
- [19] T. Bando, K. Takenaka, S. Nagasaka, and T. Taniguchi, "Unsupervised drive topic finding from driving behavioral data," in *Proc. IEEE Intelligent Vehicles Symp.*, 2013, pp. 177–182.
- [20] P. M. Fitts and J. R. Peterson, "Information capacity of discrete motor responses," *J. Exp. Psychol.*, vol. 67, no. 2, pp. 103, 1964.
- [21] D. A. Castanon, B. C. Levy, and A. S. Willsky, "Algorithms for the incorporation of predictive information in surveillance theory," *Int. J. Syst. Sci.*, vol. 16, no. 3, pp. 367–382, 1985.
- [22] Y. Bar-Shalom, P. Willett, and X. Tian, *Tracking and Data Fusion: A Handbook of Algorithms*. Bloomfield, CT: YBS Publishing, 2011.
- [23] J. Berger, *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag, 2013.
- [24] T. Carter, S. A. Seah, B. Long, B. Drinkwater, and S. Subramanian, "Ultrahaptics: Multi-point mid-air haptic feedback for touch surfaces," in *Proc. 26th Annu. ACM Symp. User Interface Software and Technology*, 2013, pp. 505–514.
- [25] Y. Zhang, S. Stellmach, A. Sellen, and A. Blake, "The costs and benefits of combining gaze and hand gestures for remote interaction," in *INTERACT 2015*. New York: Springer-Verlag, 2015, pp. 570–577.
- [26] J. R. Wilson and S. Sharples, *Evaluation of Human Work*. Boca Raton, FL: CRC, 2015.
- [27] B. I. Ahmad, J. K. Murphy, P. M. Langdon, S. J. Godsill, and R. Hardy, "Destination inference using bridging distributions," in *Proc. IEEE Acoustics Speech & Signal Processing Conf.*, Brisbane, Australia, 2015, pp. 5585–5589.
- [28] B. I. Ahmad, J. K. Murphy, P. M. Langdon, and S. J. Godsill, "Bayesian intent prediction in object tracking using bridging distributions," *IEEE Trans. Cybern.*, to be published.
- [29] A. Meucci, "Review of statistical arbitrage, cointegration, and multivariate Ornstein-Uhlenbeck," SSRN preprint 1404905, 2009.
- [30] O. Cappé, S. J. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," *Proc. IEEE*, vol. 95, no. 5, pp. 899–924, 2007.



Ahmed Badr, Ashish Khisti, Wai-Tian Tan,
and John Apostolopoulos



Perfecting Protection for Interactive Multimedia

*A survey of forward
error correction for
low-delay interactive
applications*

Many current and emerging applications require low-latency communication, including interactive voice and video communication, multiplayer gaming, multiperson augmented/virtual reality, and various Internet of Things (IoT) applications. Forward error correction (FEC) codes for low-delay interactive applications have several characteristics that distinguish them from traditional FEC. The encoding and decoding operations must process a stream of data packets in a sequential fashion. Strict latency constraints limit the use of long block lengths, interleaving, or large buffers. Furthermore, these codes must achieve fast recovery from burst losses and yet be robust to other types of loss patterns.

This tutorial article provides a survey of FEC for low-delay interactive applications. We provide several illustrative examples that explain why off-the-shelf codes such as Reed–Solomon (RS) codes, digital fountain codes, or random linear convolutional (RLC) codes do not provide ideal error correction for such applications. We then introduce some recently proposed FEC codes for streaming, discuss their properties, and quantify their performance gains through both illustrative examples and simulations over statistical channel models and real packet traces.

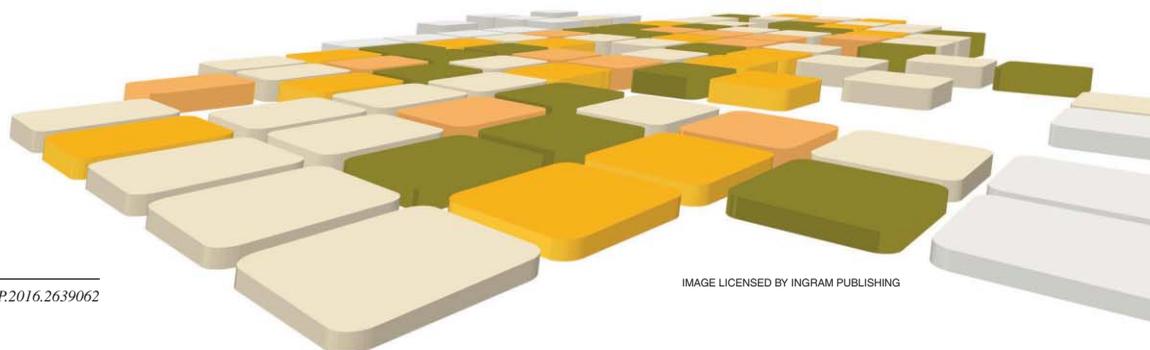


IMAGE LICENSED BY INGRAM PUBLISHING

Digital Object Identifier 10.1109/MSP.2016.2639062
Date of publication: 3 March 2017

Overview

In the last decade, we have witnessed an explosive demand for multimedia streaming applications. A recent study [1] predicts that Internet protocol (IP) video alone will constitute 79% of all consumer Internet traffic in 2018. Some commonly used applications include Voice over IP (VoIP), video on demand, videoconferencing, desktop zsharing, and interactive network gaming. Emerging applications that require low latency include augmented reality, virtual reality, and various IoT applications involving control loops for industrial processes. The underlying communication network for these applications must support high reliability, low latency, and preferably in-order delivery of source packets. Furthermore, it must include wireless links that are subjected to noise, fading, mobility, and interference. To combat such impairments, various error-control mechanisms must be implemented.

In the physical layer of wireless systems, powerful error-correcting codes, such as turbo codes, are used to combat short-term fast fading and white Gaussian noise. These codes cannot always recover from other sources of impairment, such as slow fading, buffer overflow, congestion, or interference, which cause packet losses at the application layer. It is well known that certain loss patterns, such as burst losses, can cause a significant deterioration in both audio and video streaming [2], [3]. It therefore becomes necessary to develop error-control techniques at the application layer to mitigate the effect of packet losses.

Error-control mechanisms at the application layer can be divided into two classes: error concealment and error correction. Error-concealment techniques, such as interpolation, are used to mask the effect of missing source packets. These methods are outside the scope of this tutorial. Error-correction techniques, such as retransmission and FEC, are used to achieve reliable transmission over communication links. In retransmission-based schemes—for example, automatic repeat request (ARQ)—if the transmitter receives no acknowledgment for a given packet within a certain time, the packet is retransmitted to the receiver. While retransmission is a simple and effective means of error correction, it requires point-to-point communication, a feedback channel, and low round-trip delay. The round-trip delay depends on a number of factors, such as the distance between the source and destination, the number of nodes that must be traversed in between, the processing delay at each node, and the speed of the links [4]. It is valuable to remember that even if we operate at the speed of light without any other delays, we still have end-to-end delay issues. At the speed of light, the time required to travel along the earth's circumference is 133 ms. This would correspond to the theoretical minimum round-trip delay between two diametrically opposite points on the earth's circumference. The original one-way delay of the packet transmission plus the round-trip delay from ARQ can produce a minimum latency of about 200 ms. In prac-

In augmented reality, the time from a user's motion until when it should be reflected in the user's display should be fewer than 15–20 ms to provide the experience of presence.

tice, this theoretical delay would be longer, due to the nonideal refractive index of the optical fiber and nondirect paths between the nodes. But the International Telecommunication Union recommendation states that the end-to-end latency in interactive voice and video applications must be fewer than 150 ms [5]–[7]. Clearly, even in the ideal case, the distances involved and the application constraints preclude the sole use of retransmission. Applications such as

augmented reality and virtual reality have even tighter delay constraints. For example, in augmented reality, the time from a user's motion until when it should be reflected in the user's display (commonly referred to as *motion-to-photon latency*) should be fewer than 15–20 ms to provide the experience of presence. Similarly, IoT applications that involve control feedback loops may require 1-ms or submillisecond latency, depending on the control loop requirements—orders of magnitude tighter delay constraints than traditional applications. In addition, support for ultralow-latency wireless services (millisecond level) are defined as a requirement for fifth-generation cellular systems [8].

A common alternative to retransmission is FEC, where redundant data are derived from the original data using techniques from coding theory. Error-correcting codes such as low-density parity-check and digital fountain codes [9], [10] are recommended in the Internet Engineering Task Force's real-time transport protocol profiles for noninteractive streaming applications. These codes operate over long block lengths, typically a few thousand symbols, and are thus suitable in applications in which the delay constraints are not stringent. In contrast, FEC codes used in interactive applications are often constrained to have short block lengths due to the delay constraints. Nevertheless, real-world interactive audio- and videoconferencing applications such as Skype [11] are known to use FEC with significant advantages.

In this article, we take a principled approach toward understanding ideal FEC for low-delay interactive applications. In these applications, the FEC-encoding and FEC-decoding operations must happen sequentially on the source and channel streams, respectively. Furthermore, certain erasure patterns, such as burst losses, can severely degrade the performance [2], [3], [11]. To illustrate the effect of burst losses, consider the two types of packet-loss sequences in Figure 1. Both of these sequences have the same fraction of lost packets. Sequence 1 in Figure 1 corresponds to packet losses that are well separated, while sequence 2 corresponds to packet losses in a burst. In the former case, a short-block code can be used for error correction. For example a (3, 2) RS code [12] will guarantee that all the source packets that have been erased on link 1 will be recovered. This is possible because there is, at most, one erasure in any three consecutive packets. For sequence 2 in Figure 1, a (3, 2) block code cannot be used to recover the burst loss of five packets. We will have to use a longer (15, 10) RS code to recover all the erased

source packets, while also maintaining the same overhead as in sequence 1. However, the delay incurred with this code is considerably higher than the previous case. Thus, the dynamics of packet-loss patterns, and not just the average fraction of losses, must be considered in streaming applications.

We discuss coding techniques that can repair burst losses with a much shorter delay than RS codes. We also show that codes that are optimal for burst losses in terms of minimizing the delay are rather sensitive to other loss patterns. In practice, communication links introduce both types of erasure patterns illustrated in Figure 1. Thus, we discuss coding schemes that enable fast recovery from burst losses and are also robust to isolated losses [13]–[15].

Case study: Why traditional FEC is not enough

In this section, we study the performance of various error-correcting codes in a streaming setup via an example. To provide a common point of comparison, we focus on the streaming setup shown in Figure 2. In this model, a source packet $\mathbf{s}[t]$ for $t = 0, 1, 2, \dots$ arrives at the FEC encoder every t_s seconds—i.e., $\mathbf{s}[t]$ arrives at time $t \cdot t_s$ seconds. For simplicity, we will assume that each source packet is of the same size and consists of k symbols. The encoder generates a channel packet $\mathbf{x}[t]$ of size n symbols and transmits it in the interval $[t \cdot t_s, (t + 1) \cdot t_s)$. The encoding function is causal

$$\mathbf{x}[t] = f_i(\mathbf{s}[t - m], \dots, \mathbf{s}[t]), \quad t \geq 0, \quad (1)$$

where $f_i(\cdot)$ is the encoding function at time t and m denotes the memory of the encoder. Furthermore, the rate of the code is given by $R = k/n$, and its redundancy is $100(n - k)/k\%$.

The dynamics of packet-loss patterns, and not just the average fraction of losses, must be considered in streaming applications.

The communication channel considered is a packet erasure channel. Each transmitted packet is either erased or perfectly received at the destination. This is motivated by the fact that erroneous packets are discarded at lower layers in the communication protocol stack. In particular, the channel output at (discrete) time t is given by $\mathbf{y}[t] = \star$ if the

channel introduces an erasure at time t , and by $\mathbf{y}[t] = \mathbf{x}[t]$ if it does not. Throughout this article, we will use the term *channel* to denote the packet-loss sequence, as is the convention in the coding theory literature. To develop insights into the performance of different coding schemes, we focus on a simple class of channels defined below.

Definition 1 (burst-erasure channel)

A burst-erasure channel with parameter B is a channel that introduces a single contiguous sequence of erasures of maximum length B . That is, starting from some arbitrary time $j \geq 0$ and $0 \leq B' \leq B$, we have that $\mathbf{y}[t] = \star$ for $t \in [j, j + B' - 1]$ and that $\mathbf{y}[t] = \mathbf{x}[t]$ otherwise.

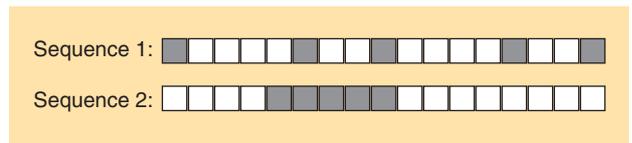


FIGURE 1. Two examples of erasure sequences that have the same number of erasures but different erasure patterns. The shaded boxes denote the erasures, while the white boxes denote packet reception. In sequence 1, the erasures are mostly isolated, while in sequence 2, they occur in a single burst. One can use a short (3, 2) RS code to recover for sequence 1, but a longer (15, 10) RS code is required over sequence 2, resulting in a higher delay.

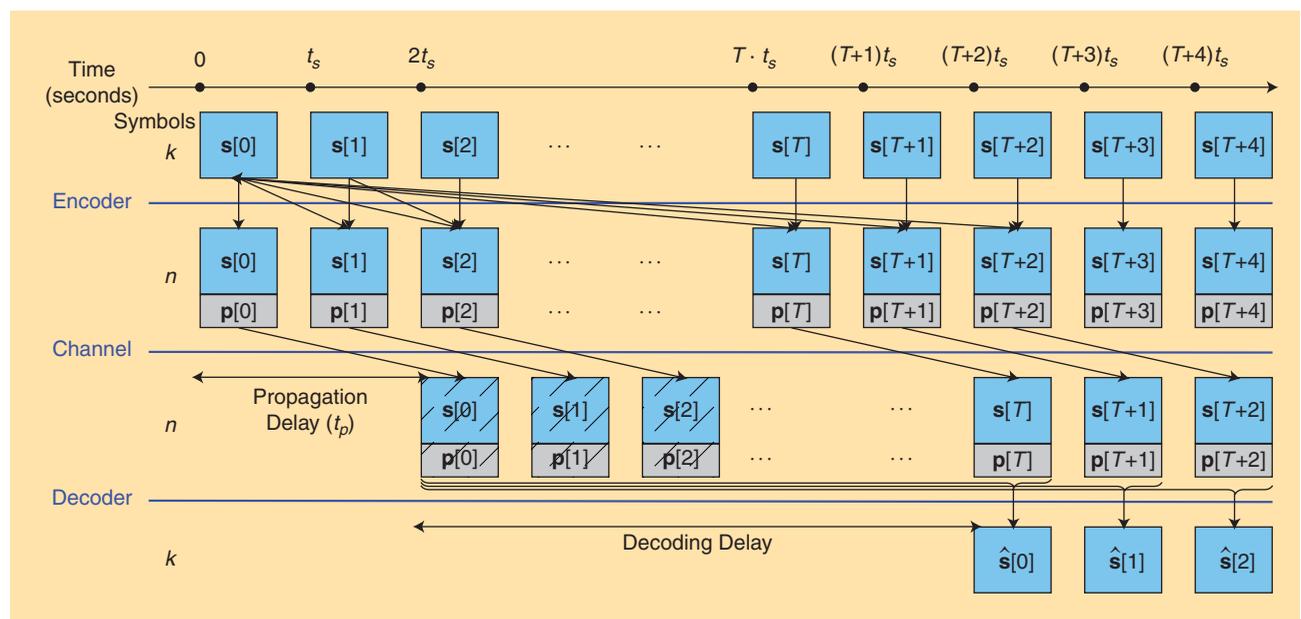


FIGURE 2. The source stream $\mathbf{s}[t]$ for $t \geq 0$ is encoded to a channel stream $\mathbf{x}[t]$, which is transmitted over an erasure channel. The decoder tolerates a maximum decoding delay of T packets.

Definition 2 (isolated-erasure channel)

An isolated-erasure channel with parameter N is a channel that introduces up to N erasures in the received stream. The location of the erasures can be arbitrary. Thus, for some $0 < N' \leq N$ and $0 \leq j_1 < j_2 \dots < j_{N'}$, we have that $\mathbf{y}[j_i] = *$, and $\mathbf{y}[t] = \mathbf{x}[t]$ if $t \notin \{j_1, j_2, \dots, j_{N'}\}$.

The channel models treated above are rather simple—the burst erasure model introduces a single burst of maximum length B , while the isolated-erasure channel introduces a maximum of N erasures in arbitrary locations. Nevertheless, there are several advantages in studying these models:

- 1) The study of such simplified models provides first-order insights into the performance of various streaming codes. For example, we will see how convolutional codes are more resilient than block codes in the streaming setup.
- 2) The analysis of these channels is a useful first step in treating more sophisticated models, such as the sliding window channel models, which must be naturally considered in streaming scenarios [13], [14].
- 3) We will see that the insights obtained through the study of such channels will be useful in interpreting the simulation results over the Gilbert–Elliot model and real packet traces treated in the “Simulation Results” section.

As shown in Figure 2, the decoder tolerates a maximum delay of T packets. That is,

$$\hat{\mathbf{s}}[t] = \gamma_t(\mathbf{y}[0], \dots, \mathbf{y}[t + T]), \quad (2)$$

where $\gamma_t(\cdot)$ designates the decoding function at time t . The source packet $\mathbf{s}[t]$ is declared lost if $\hat{\mathbf{s}}[t] \neq \mathbf{s}[t]$. A delay of T

packets in our model is equivalent to an actual delay of $(T \cdot t_s + t_p)$ seconds, where t_s is the interpacket arrival time and t_p is the propagation delay in Figure 2. In the rest of this article, we consider the delay in terms of packets, and the time index will refer to the discrete time.

Remark 1

The constructions considered in Figure 2 are systematic codes, i.e., each channel packet can be expressed as $\mathbf{x}[t] = (\mathbf{s}[t], \mathbf{p}[t])$, where $\mathbf{p}[t]$ is the parity-check packet consisting of $(n - k)$ symbols. All codes that we consider in this article are systematic codes. This will guarantee that whenever a channel packet is received, the underlying source packet is immediately recovered, with zero delay. Furthermore, all codes we consider are linear codes, i.e., the parity-check symbols can be expressed as a linear combination of the source packets [16].

Remark 2

In the setup in Figure 2, the parity-check packets $\mathbf{p}[t]$ are not transmitted as separate packets but are appended to the source packets before transmission. This reduces the number of packets transmitted over the channel. Such an approach is desirable in practical wireless networks such as 802.11, where channel contention overhead is significant. Nevertheless, most of the insights developed for our proposed model also apply, with minor variations, to the case in which the parity-check packets are transmitted separately. The advantages of using separate FEC streams include wider compatibility, where the media stream can be decoded even by clients that do not understand FEC.

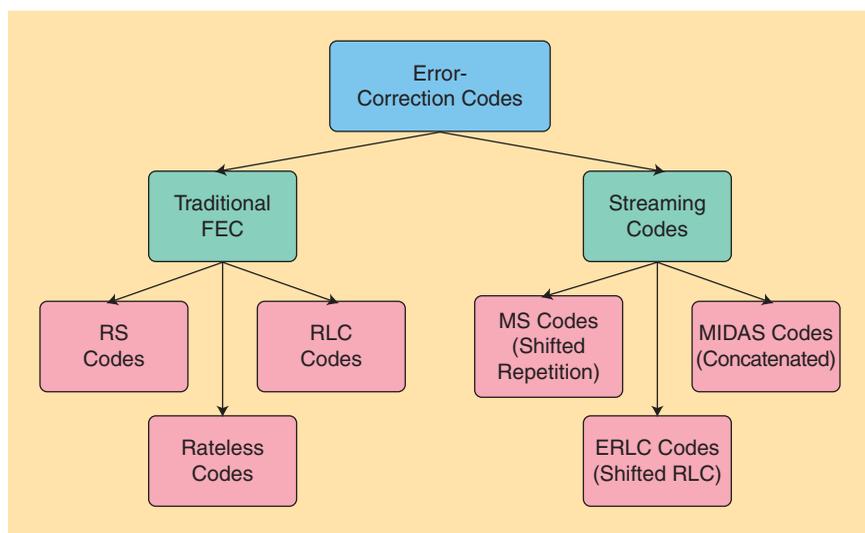


FIGURE 3. A summary of different coding schemes in the streaming setup. The traditional FEC codes are discussed in the “Case Study” section. The rate-1/2 streaming codes—shifted-repetition, shifted-RLC, and concatenated codes—are also discussed in the “Case Study” section. The shifted-repetition code provides optimal burst-error correction in the streaming setup, while the shifted-RLC and concatenated codes are a robust extension of these codes. Their respective generalizations—MS codes, ERLC codes, and MIDAS codes—are discussed in the “General Code Constructions” section.

Summary of coding schemes

We briefly summarize the different code constructions discussed in this article. As illustrated in Figure 3, the coding schemes we consider can be broadly classified into two categories: 1) traditional FEC and 2) streaming codes. In the former category, we discuss three off-the-shelf coding schemes: RS codes, rateless codes, and RLC codes in the next three sections, respectively. A common feature of these codes is that, following a loss pattern, the decoder must collect enough parity checks so that it can invert the resulting system of equations and simultaneously recover all the missing source packets. For example, when the rate of the code is $R = 1/2$, so that the size of each parity check equals that of the source packet, the decoder must collect as many parity-check packets as missing source

packets to recover them. In the special case of the burst-erasure channel with burst length of B and $R = 1/2$, this results in a delay of $T = 2B$.

One can significantly improve upon the performance of traditional FEC over burst-erasure channels. Such constructions are referred to as *streaming codes* in this article. Unlike traditional FEC, they do not force simultaneous recovery of all the source packets. Instead, the construction of parity checks is such that the older source packets with earlier deadlines are recovered before the later source packets. The streaming codes discussed in this article are illustrated in Figure 3. The minimum delay achieved by this method is $T = B$, when $R = 1/2$.

In Figure 4, we provide a comparison between traditional FEC and streaming codes. We sketch the maximum correctable burst length on the x axis and the resulting delay for different codes on the y axis. The rate of all codes is fixed to $R = 1/2$. As we discussed, when the burst length equals B , the minimum delay for traditional FEC is $T = 2B$, which is shown by the blue line in the figure. The associated region $T \geq 2B$ is shaded light blue. In contrast, the minimum delay achieved by streaming codes is $T = B$ and is shown by the red line. Thus, the longer the burst length, the higher will be the gain provided by streaming codes. As we will see, the codes achieving minimum delay over burst loss channels are sensitive to other erasure patterns. Thus, in practice one must develop robust extensions that are also resilient to isolated erasure patterns. Such codes will require slightly larger delays than $T = B$ and will achieve a performance in the light-red region shown in Figure 4. We discuss three such constructions—shifted RLC codes, concatenated codes, and dual-delay codes—in the respective eponymous sections below. The corresponding generalizations to arbitrary rates are discussed in the “General Code Constructions” section.

RS codes

An (n, k) block code operates on k source packets and generates $n > k$ packets. Hence, the rate of an (n, k) code is given by k/n . Systematic codes are a class of block codes where the first k packets of the codeword are the source packets, whereas the last $n - k$ are parity-check packets. RS codes [12], [16] are the most commonly used block codes. These codes belong to the class of maximum distance separable (MDS) codes, which guarantee the recovery of the maximum number of packet losses for a given redundancy. An (n, k) RS code can recover up to $n - k$ erased packets in any codeword of length n .

While an (n, k) block code does not directly fit into the streaming setup, it can be easily adapted, as discussed below. The stream of source packets is logically split into segments, each of size k . An (n, k) block code is then applied to each segment to generate $n - k$ parity-check packets. These parity-check packets are then transmitted together with the source packets in the next block of k packets. This construction is particularly simple for $R = 1/2$, which is the case treated in this section. For the case of general rates, we refer the reader to [17].

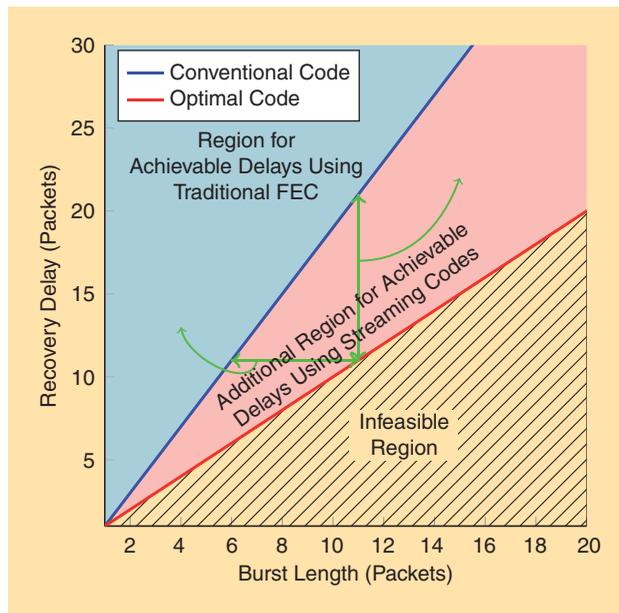


FIGURE 4. The achievable delays for erasure recovery of different burst lengths using FEC at rate 1/2. The solid red line shows the minimum delay that can be achieved for a given burst length. The delay below this threshold cannot be achieved by any code. The blue region shows the delay achieved by traditional FEC codes that perform simultaneous recovery of the source packets. Streaming codes that perform sequential recovery can achieve delay in the red region. By tracing the regions marked by the green line, we can see that at a recovery delay of $T = 11$ and $R = 0.5$, it is possible to achieve an optimal burst length recovery of 11 packets versus the conventional approach, which can only recover six packets. Conversely, for recovering a burst length of 11, the traditional codes require a delay of 21 packets while the optimal code only requires a delay of 11 packets.

In Figure 5, a $(4, 2)$ RS code is applied to each group of two consecutive source packets to generate two parity-check packets. For example, in the first block we generate

$$(\mathbf{s}[0], \mathbf{s}[1]) \xrightarrow{(4,2) \text{ RS Code}} (\mathbf{s}[0], \mathbf{s}[1], \mathbf{p}[2], \mathbf{p}[3]). \quad (3)$$

The resulting parity-check packets $(\mathbf{p}[2], \mathbf{p}[3])$ are transmitted in the next block by appending them to $\mathbf{s}[2]$ and $\mathbf{s}[3]$, respectively. The resulting channel packets are $\mathbf{x}[2] = (\mathbf{s}[2], \mathbf{p}[2])$ and $\mathbf{x}[3] = (\mathbf{s}[3], \mathbf{p}[3])$. More generally, for the group of source packets $(\mathbf{s}[2i], \mathbf{s}[2i + 1])$, a $(4, 2)$ RS code is applied to generate parity checks $\mathbf{p}[2i + 2]$ and $\mathbf{p}[2i + 3]$, which are transmitted along with the source packets at times $t = 2i + 2$ and $t = 2i + 3$, respectively.

A longer $(6, 3)$ RS code can be applied in an analogous fashion by considering groups of three source packets and generating three parity-check packets, which must be transmitted in the next block of three source packets.

We now discuss the error-correction properties based on three following cases.

- 1) *Single isolated loss:* Consider a channel that introduces a single isolated erasure, i.e., $N = 1$. The $(4, 2)$ RS code can recover the missing source packet with a delay of no more than $T = 2$ packets. For example, if $\mathbf{x}[0]$ is lost, then the associated source packet is recovered as soon as $\mathbf{p}[2]$ is

received by the decoder. In contrast, the (6, 3) RS code can recover the missing source packet with a worst-case delay of $T = 3$ packets.

- 2) *Two isolated losses*: Next, consider the case in which the channel introduces up to two isolated losses. For the (4, 2) code, it can be seen that the worst-case delay happens when $\mathbf{x}[0]$ and one of either $\mathbf{x}[1]$ or $\mathbf{x}[2]$ is erased. The source packet $\mathbf{s}[0]$ can be recovered from $\mathbf{p}[3]$, resulting in a delay of $T = 3$ packets. Similarly for the (6, 3) code, the worst-case delay with two isolated losses is $T = 4$. This will happen, for example, if $\mathbf{x}[0]$ and $\mathbf{x}[3]$ are erased, in which case the decoder must wait for $\mathbf{p}[4]$ to recover $\mathbf{s}[0]$.
- 3) *Burst-erasure channel*: Finally, consider the case in which the channel introduces a burst of length $B = 3$. In particular, suppose that $\mathbf{x}[0]$, $\mathbf{x}[1]$, and $\mathbf{x}[2]$ are erased. The (4, 2) RS code will not be able to recover $\mathbf{s}[0]$ and $\mathbf{s}[1]$, although $\mathbf{s}[2]$ can still be recovered from $\mathbf{p}[4]$. In contrast, the (6, 3) RS code successfully recovers all the erased source packets with a maximum delay of $T = 5$.

Generally speaking, longer block codes in the streaming setup will correct from longer bursts but at the expense of longer delay. However, the size of each block must be small because of the delay constraints. Such an approach significantly limits the error-correction capability. As we will see, the use of convolutional codes is more desirable than block codes, as it enables the decoder to recover from shorter bursts with smaller delays, while longer bursts can be recovered with longer delays. However, before discussing these, we briefly discuss rateless codes.

Rateless codes

RS codes exist over fields of sizes at least as large as the block length. Typical block lengths for RS codes are restricted to $n \leq 255$. Rateless codes (e.g., Luby transform codes [9] and raptor codes [10]) are a class of binary codes that can support considerably longer block lengths that achieve near-optimal error correction and are amenable to extremely efficient decoding algorithms. This makes them a natural choice in

The use of convolutional codes is more desirable than block codes, as it enables the decoder to recover from shorter bursts with smaller delays.

noninteractive streaming applications. However, since the focus of this article is on FEC for interactive applications, rateless codes will not be suitable.

RLC codes

Together with block codes and rateless codes, convolutional codes [16], [18] form a commonly implemented class of error-correcting codes. Such codes have an inherent sequential encoding structure. At each time instant t , an (n, k, m) convolutional code generates one channel packet $\mathbf{x}[t]$ of size n that is a causal combination of the previous m source packets and the current packet—i.e., $\mathbf{x}[t] = f_i(\mathbf{s}[t-m], \dots, \mathbf{s}[t-1], \mathbf{s}[t])$. The rate of such a code is given by $R = k/n$, and its redundancy is $100(n-k)/k\%$. The code is said to be systematic if each channel packet $\mathbf{x}[t]$ contains the source packet $\mathbf{s}[t]$ —i.e., $\mathbf{x}[t] = (\mathbf{s}[t], \mathbf{p}[t])$, where $\mathbf{p}[t]$ is the size $n-k$ parity-check packet at time t . An important class of convolutional codes is linear, time-invariant convolutional codes, where the parity-check packets can be expressed as

$$\mathbf{p}[t] = \sum_{i=1}^m \mathbf{s}[t-i] \cdot \mathbf{H}_i, \quad (4)$$

where m denotes the memory of the code and the matrices \mathbf{H}_i are of dimension $k \times n-k$ for each $i = 1, \dots, m$. The summation in (4) starts at $i = 1$ and not $i = 0$, i.e., $\mathbf{p}[t]$ does not combine $\mathbf{s}[t]$, because the packet erasure channel considered will erase $\mathbf{p}[t]$ whenever $\mathbf{s}[t]$ is erased. Furthermore, when $\mathbf{x}[t]$ is not erased, the systematic code will recover $\mathbf{s}[t]$ directly without the need of $\mathbf{p}[t]$.

If the coefficients in the matrix \mathbf{H}_i are selected at random, then the codes are said to be *RLC codes* (see, e.g., [19] and [20]). Such codes guarantee that, with high probability, each parity-check symbol provides an independent equation involving the source symbols. One can also construct the matrices \mathbf{H}_i in a deterministic fashion to satisfy this property. Such constructions also achieve the largest distance up to the code memory and are referred to as *strongly MDS codes* (see, e.g., [21] and [22]). For simplicity, we will refer to all these constructions as *RLC codes*.

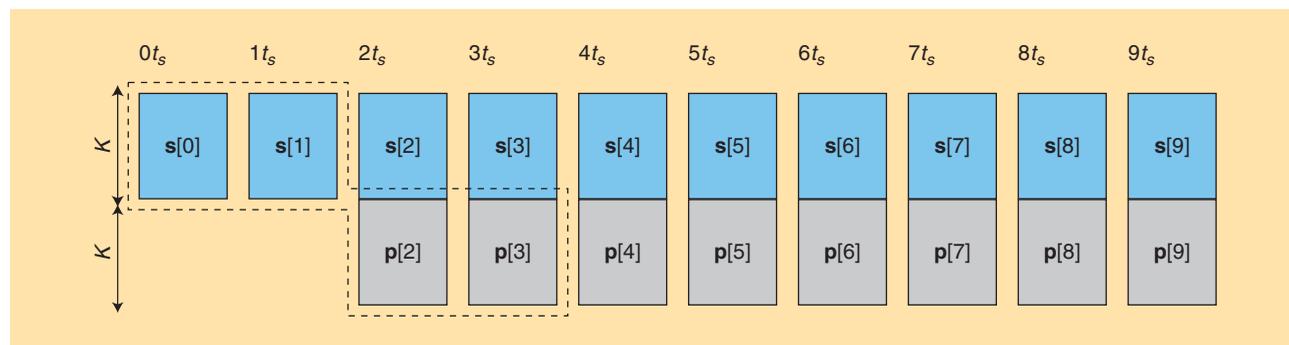


FIGURE 5. A (4, 2) RS code applied to a streaming setup. The parity-check packets $\mathbf{p}[2]$ and $\mathbf{p}[3]$ are generated from $\mathbf{s}[0]$ and $\mathbf{s}[1]$ but are sent in the next block.

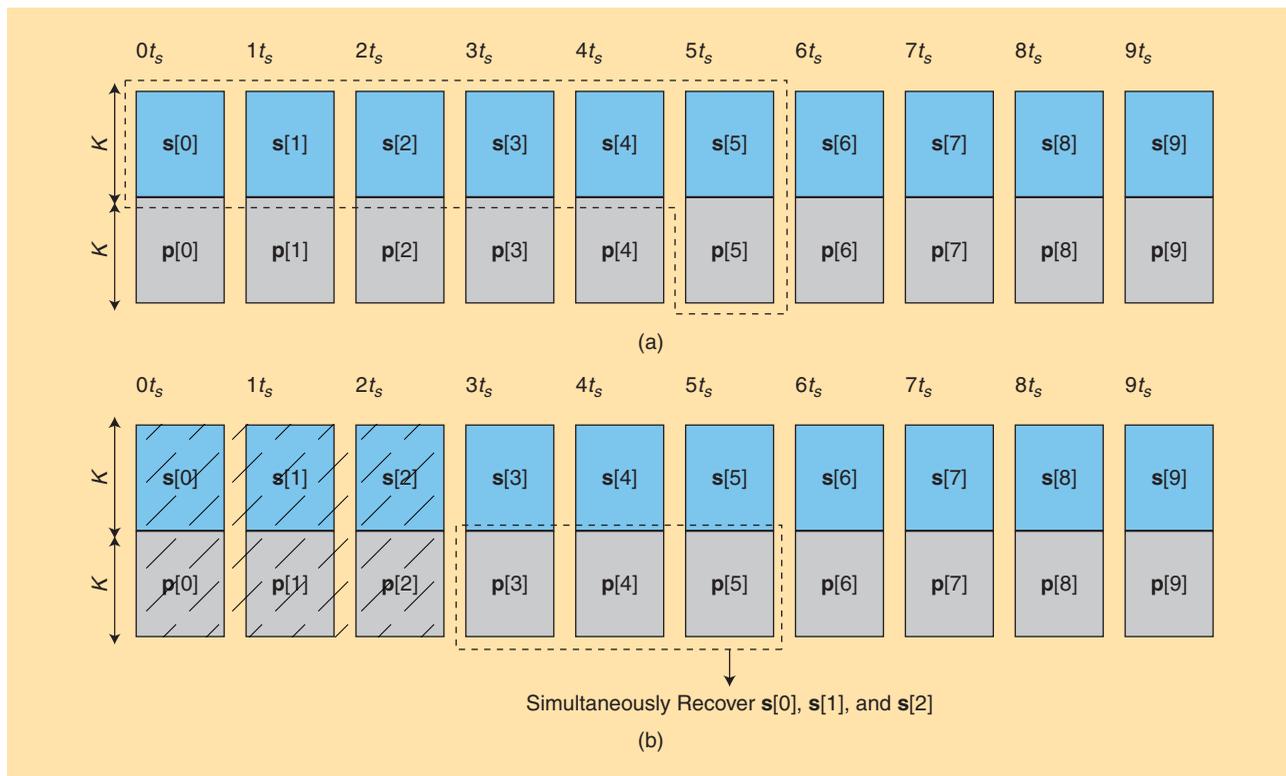


FIGURE 6. A (2, 1, 5) RLC code: (a) encoder and (b) burst of length three.

Figure 6 illustrates a (2, 1, 5) RLC code of rate 1/2. In this special case, the parity-check packets are the same size as the source packets. We can express this as

$$p[t] = \sum_{i=1}^5 \alpha_i \cdot s[t - i], \quad (5)$$

where α_i are scalars instead of the matrices in (4). We analyze the performance of these codes for the same set of erasure patterns as in the case of block codes. A convenient shorthand that we adopt for the rest of the article is to omit explicit representation of the coefficients α_i and simply express the parity as sum of source symbols, implicitly assuming that each symbol is also multiplied by a non-zero α_i .

- 1) *Single isolated loss*: Consider a channel that introduces a single isolated erasure, i.e., $N = 1$. In particular, suppose that $x[0]$ is lost. It is clear that the erased source packet $s[0]$ is recovered as soon as the channel packet $x[1]$ —and in particular $p[1]$ —is obtained, i.e., with a delay of $T = 1$ packet. Thus, the RLC code achieves a smaller delay than the (4, 2) and (6, 3) RS codes.
- 2) *Two isolated losses*: Next, consider the case in which the channel introduces up to two isolated losses. It can be verified that the worst-case delay occurs when the two losses happen in succession—for example, if $x[0]$ and $x[1]$ are erased. In this case, both source packets $s[0]$ and $s[1]$ are recovered when $p[2]$ and $p[3]$ are received, i.e., with a delay of $T = 3$. This is the same delay as the shorter (4, 2) RS code.

- 3) *Burst-erasure channel*: Finally, consider the case in which the channel introduces a burst of length $B = 3$, and in particular suppose that $x[0]$, $x[1]$, and $x[2]$ are erased. The RLC code will collect the parity-check packets $p[3]$, $p[4]$, $p[5]$ and then recover all the erased source packets— $s[0]$, $s[1]$, $s[2]$ —simultaneously, with a delay of $T = 5$. This is illustrated in Figure 6(b). This is the same delay as the (6, 3) RS code. Furthermore, since the memory $m = 5$, the decoder can also recover a burst of length $B = 4$ with a delay of $T = 7$ and a burst of length $B = 5$ with a delay of $T = 9$. These patterns cannot be corrected by the RS codes previously discussed.

Based on the above discussion, it is clear that convolutional codes exhibit several advantages over block codes. We summarize these below.

- 1) Unlike block codes, convolutional codes do not require the source sequence to be fragmented into blocks over which the parity checks are generated. Instead, they are based on a sliding-window construction [compare (4)]. This approach enables the decoder to opportunistically recover shorter burst lengths more quickly than longer bursts, as we discussed in the previous example.
- 2) The memory of the code m is a design parameter. Larger values of m will enable longer burst lengths to be recovered at the same rate. However, longer memory increases complexity and also makes the code vulnerable to certain other types of erasure patterns when partial recovery is the best option. To explain this, consider a rate $R = 1/2$ RLC with infinite memory, and one with memory $m = 5$. Suppose the

channel introduces a burst of length $B = 20$ in the interval $t \in [i, i + 19]$. The infinite-memory code will force the decoder to use the next 20 parity checks in the interval $[i + 20, i + 39]$ to recover the erased source sequence. Any additional losses in this period will cause longer delays. The code with memory $m = 5$ will behave very differently. It will skip parity checks in the interval $[i + 20, i + 24]$, which are the only received parities that depend on the burst interval. Thereafter, any parity checks can be used to recover from any future losses. Thus, due to delay constraints, the code with memory $m = 5$ is more desirable in the event of such burst losses.

It should be noted that the construction in (4) applies to any arbitrary rate R . There is nothing special about $R = 1/2$, except the simple construction (5). The following result shows the burst- and isolated-error correction properties of RLC codes [13] for an arbitrary rate R .

Theorem 1 (error-correction properties of RLC codes at a given maximum delay)

Consider an (n, k, m) RLC code with rate $R = (k/n)$ and memory $m \geq T$. Such a code can recover from a burst-erasure channel with a maximum burst length B , or from an isolated-erasure channel with a maximum of N erasures, with a maximum delay of T , provided that

$$B \leq (1 - R)(T + 1), \quad (6)$$

$$N \leq (1 - R)(T + 1). \quad (7)$$

RLC codes have the same threshold for burst-error and isolated-error correction. To explain this, recall that RLC codes perform simultaneous recovery of the source packets in the event of an erasure burst. They treat each parity check as providing an equation involving the source symbols and recover all the erased symbols simultaneously when sufficiently many parity checks are received. This is illustrated in Figure 6(b). They are not able to recover in an opportunistic fashion earlier source packets whose deadlines occur earlier. In the following four sections, we discuss the class of streaming codes that can achieve such a sequential recovery and thus provide improved performance over burst-erasure channels.

Shifted-repetition code

A repetition code is a simple construction with rate $R = 1/2$, where each source packet is repeated with a unit delay—i.e., $\mathbf{x}[i] = (\mathbf{s}[i], \mathbf{s}[i - 1])$ for all $i \geq 1$. While simple in implementation, such a construction cannot recover from burst losses of length $B \geq 2$. Interestingly, a simple variation of this construction achieves optimal recovery over the burst-erasure channel. Some generalizations of repetition codes, where low-bit-rate redundant audio packets are used as parities, are studied in [23].

In contrast to RLC codes, the parity-check packets in a shifted-repetition code do not involve a linear combination of the source packets.

A shifted-repetition code is a rate $R = 1/2$ code, where each source packet is repeated once after a delay of T packets—i.e., we can express $\mathbf{x}[i] = (\mathbf{s}[i], \mathbf{s}[i - T])$. In contrast to RLC codes, the parity-check packets in a shifted-repetition code do not involve a linear combination of the source packets. We replace (4) with simply

$\mathbf{p}[i] = \mathbf{s}[i - 5]$. We note the following properties:

- 1) *Single isolated loss*: When there is a single isolated loss, the corresponding source packet can be recovered with a delay of $T = 5$ packets. For example, if $\mathbf{x}[0]$ is lost, then the source packet $\mathbf{s}[0]$ is recovered when its repeated copy at time $T = 5$ is received.
- 2) *Two isolated losses*: In general, the shifted-repetition code cannot recover from two or more isolated losses. As an example, if the erasures happen at time $t = 0$ and $t = 5$, then the source packet $\mathbf{s}[0]$ cannot be recovered. Thus, the delay for this case is ∞ .
- 3) *Burst-erasure channel*: The shifted-repetition code can correct a burst of length $B = 5$ with a delay of $T = 5$. Suppose that the erasure burst spans the interval $[0, 4]$. Then $\mathbf{s}[0]$ is recovered at time $t = 5$ from $\mathbf{p}[5] = \mathbf{s}[0]$. Likewise, each $\mathbf{s}[j]$ for $j = 0, \dots, 4$ is recovered at time $t = j + 5$ in a sequential manner.

It is clear that a shifted-repetition code with delay T will recover any burst of length $B \leq T$. This is clearly the maximum burst length that can be recovered by any code. However, the rate of the code is fixed at $R = 1/2$. Maximally short (MS) codes [24], [25] are a generalization of the shifted-repetition code that achieve optimal burst correction. For a given rate R and delay T , they achieve $B = \min(1, (1 - R)/R)T$. We review a variation of the original construction in the “General Code Constructions” section. It should be noted that the value of B is larger than that of RLC codes in Theorem 1. Unfortunately, like the shifted-repetition codes, these codes are sensitive to the isolated-erasure channel with $N \geq 2$. We will see that this can lead to a significant degradation over statistical channels, such as the Gilbert–Elliott channel. Nevertheless, the MS codes constitute an important building block for the more robust codes discussed in the sequel.

Shifted-RLC codes

Shifted-RLC codes combine concepts of shifted-repetition code with RLC code. They achieve a longer burst-error correction threshold than RLC codes in Theorem 1, but smaller than the shifted-repetition codes. However, unlike the shifted-repetition codes, they can correct from more than one isolated loss. As an example, consider the rate-1/2 code $\mathbf{x}[i] = (\mathbf{s}[i], \mathbf{p}[i])$, where we select

$$\mathbf{p}[i] = \mathbf{s}[i - 5] + \mathbf{s}[i - 4].$$

This code is similar to the $(n = 2, k = 1, m = 2)$ RLC code, but the parity-check packets are further delayed by $\Delta = 3$

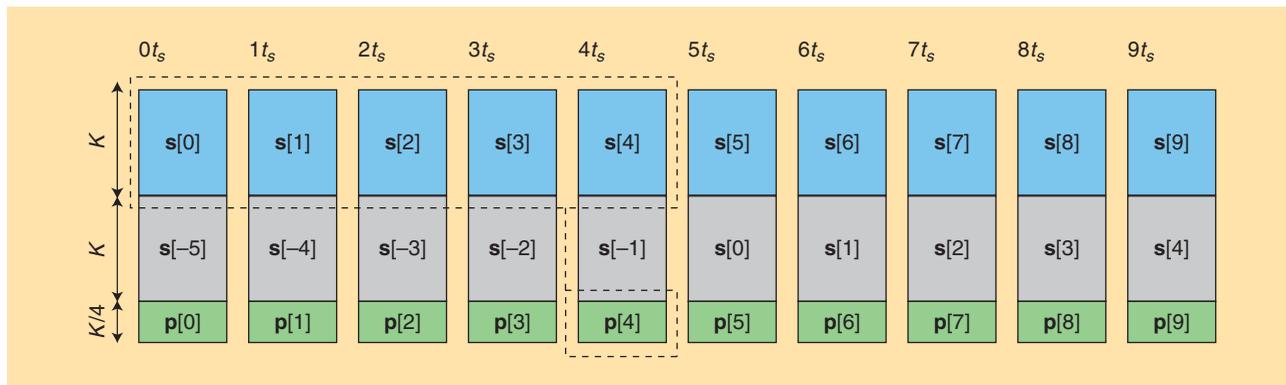


FIGURE 7. A rate-4/9 concatenated code.

units. To reiterate, in general, there will be nonzero coefficients multiplying the source symbols $s[i - 5]$ and $s[i - 4]$. These have been suppressed as discussed previously. We summarize the error correction properties as follows:

- **Single isolated loss:** When there is a single isolated loss, the corresponding source packet can be recovered with a delay of $T = 4$ packets. For example, if $x[0]$ is erased, $s[0]$ is recovered when $p[4] = s[0] + s[-1]$ is available to the decoder.
- **Two isolated losses:** This code recovers from any pattern consisting of $N = 2$ erasures within a worst-case delay of $T = 5$. The worst-case pattern corresponds to $x[0]$ and $x[4]$ being erased. For this pattern, the first available parity check that involves $s[0]$ is at time $T = 5$. Using $p[5] = s[1] + s[0]$, the source packet $s[0]$ is recovered at $T = 5$.
- **Burst-erasure channel:** A burst of length $B \leq 4$ packets is recoverable with a delay of $T = 4$. For example, suppose that $x[0], \dots, x[3]$ are erased. Then using $p[4] = s[0] + s[-1]$ and canceling $s[-1]$, which is not erased, the decoder can recover $s[0]$. Similarly, at time $t = 5$, the decoder can use $p[5] = s[0] + s[1]$ to recover $s[1]$. Continuing this process, each erased packet is recovered sequentially with delay $T = 4$.

The shifted-RLC code corrects a maximum burst length of $B = 4$ and up to $N = 2$ isolated losses within a worst-case delay of $T = 5$. For the same delay of $T = 5$, the shifted-repetition code recovers a burst length of $B = 5$, while the RLC code in Theorem 1 can recover from a burst length of $B = 3$, as well as $N = 3$ isolated losses. The main design parameter in the shifted-RLC code is the shift Δ applied to the parity-check packets. Selecting $\Delta = 0$, we recover the original RLC construction, which results in the error-correction thresholds stated in Theorem 1. Selecting $\Delta = T$ will result in the same performance as the shifted-repetition code. By selecting the value of Δ in between these two extremes, we can trade off the burst-error and isolated-error correction capabilities of the code.

Embedded RLC (ERLC) is a generalization of the aforementioned shifted-RLC code to arbitrary rates [15]. In this

construction, too, a graceful tradeoff between the burst-error correction and the isolated-error correction capabilities can be obtained through the choice of the shift parameter Δ . These codes are reviewed in the “General Code Constructions” section.

Concatenated codes

An alternative technique for making shifted-repetition codes resilient to the isolated-erasure channel model is to append an extra layer of parity checks. In Figure 7, we illustrate a concatenated code of rate $R = 4/9$, which combines a shifted-repetition code and an RLC code. The encoding steps are as follows:

- 1) We construct a rate-1/2 shifted-repetition code with a delay of $T = 5$. Each source packet $s[i]$ is repeated with a delay of $T = 5$ as shown.
- 2) We apply an $(n = 5k/4, k, m = 5)$ RLC code to the source packet $s[i]$ to generate parity-check packets of size $k/4$. These parity-check packets are appended to the source packets to generate the channel packet: $x[i] = (s[i], s[i - 5], p[i])$.

The rate of the above construction is $4/9$, which is lower than other codes discussed in this section. The construction of the rate-1/2 code in this family is a little more complicated. The construction for general rates will be discussed in the subsequent section. Nevertheless, this code is effective against burst and isolated erasures as discussed next.

- 1) **Single isolated loss:** When there is a single isolated loss, the corresponding source packet can be recovered with a delay of $T = 4$ packets. For example, if $x[0]$ is erased, $s[0]$ is recovered when $p[1], \dots, p[4]$ become available using the RLC code. Alternatively, the repetition code can also be used to recover $s[0]$, albeit with a delay of $T = 5$ packets.
- 2) **Two isolated losses:** This code recovers from $N = 2$ isolated erasures within a worst-case delay $T = 5$. The worst-case pattern corresponds to an erasure at $t = 0$ and an additional erasure in the interval $[1, 4]$. This will force the decoder to use the repetition code to recover $s[0]$, resulting in a delay of $T = 5$.

3) *Burst-erasure channel*: A burst of length $B \leq 5$ packets is recoverable with a delay of $T = 5$ by simply using the shifted-repetition constituent code and ignoring the RLC code.

For the burst-erasure channels, the concatenated code above recovers from the same burst lengths as the shifted-repetition code. For the isolated-erasure channel, it has the same performance as the shifted RLC. However, the rate of this code is $R = 4/9$, which is smaller than the other codes that achieve $R = 1/2$. A generalization of this approach to arbitrary rates, known as the maximum distance and span (MIDAS) code, is introduced in [13] and [14]. Similarly, this construction is obtained through an extension of MS codes by appending an extra layer of RLC parity checks, as explained in the “General Code Constructions” section.

Dual-delay codes

Although the shifted-RLC and concatenated codes in the latter two sections can recover from both isolated and burst erasures, they incur long delays even in the case of a single erasure. In this section, we discuss another construction that quickly recovers from a single erasure while keeping good burst correction capabilities. The rate-1/2 version of such codes is a simple combination of two shifted-repetition codes with delays of 1 and 5. That is, the parity-check packet at time i is given by

$$\mathbf{p}[i] = \mathbf{s}[i - 1] + \mathbf{s}[i - 5].$$

The achievable recovery delays for different erasure patterns are as follows:

- *Single isolated loss*: In case of a single erasure, the corresponding packet can be recovered with a delay of $T = 1$

packet. For example, if $\mathbf{x}[0]$ is erased, $\mathbf{s}[0]$ is recovered at time $t = 1$ since $\mathbf{p}[1] = \mathbf{s}[0] + \mathbf{s}[-4]$.

- *Two isolated losses*: This code recovers from $N = 2$ isolated erasures within a worst-case delay $T = 5$. The worst-case pattern corresponds to two consecutive erasures at time $t = 0$ and $t = 1$. The decoder has to wait until $\mathbf{p}[5] = \mathbf{s}[4] + \mathbf{s}[0]$ becomes available.
- *Burst-erasure channel*: A burst of length $B = 4$ packets is recoverable with a delay of $T = 5$. In this case, the parities $\mathbf{p}[5], \dots, \mathbf{p}[8]$ can be used to recover the erased $\mathbf{s}[0], \dots, \mathbf{s}[3]$, respectively.

While we do not discuss these codes in further detail in this article, the interested reader is referred to [51].

Numerical comparisons

Table 1 summarizes the properties of various error-correction codes discussed in the previous section. We set the worst-case delay of each code to be at most $T = 5$ and find the maximum burst length that can be corrected by each. All codes except the concatenated code have a rate of $R = 1/2$. The rate of the concatenated code is $R = 4/9$. Among all the codes, the shifted-repetition code achieves the maximum value of $B = 5$. However, it cannot recover from the isolated-erasure channel with $N \geq 2$. For such a channel, the RLC codes clearly outperform all other codes. However, they can only correct a burst of length $B = 3$. The shifted-RLC code and the concatenated code achieve $B = 4$ and are feasible against isolated erasures, albeit with higher delays than RLC codes. In the “Simulation Results” section, we further compare the performance of these codes at $T = 5$ over Gilbert–Elliott channels and real packet traces.

Impact on applications

As noted before, the maximum allowable one-way latency in interactive applications should not exceed 150 ms. In a VoIP application where each audio packet spans 10–20 ms of speech and assuming a 30–40-ms propagation delay for coast-to-coast communication [4], this corresponds to a maximum allowed delay of $T = 5–12$ packets. In a typical video application at 2 Mb/s and packet sizes of 1,500 B, a delay of $150 - 30 = 120$ ms will be $T \approx 20$ packets [15].

The codes in Table 1 can be naturally extended to recover from an arbitrary burst length B and delay T . Figure 8 provides an extension of Figure 4 to include the robust extensions. The uppermost black line corresponds to RLC codes, while the lowermost red line corresponds to shifted-repetition (MS) codes as before. Codes such as shifted-RLC (ERLC) and concatenated codes (MIDAS) require a slightly larger delay for burst-error correction but are also robust to isolated losses. The rate is set to be $R = 1/2$ for all codes in Figure 8. As an example, consider $B = 11$. We observe that the delay achieved by an RLC code equals $T = 21$, while the shifted-repetition (MS) code achieves $T = 11$. The two robust extensions, which can

Table 1. A summary of the error correction codes discussed in the “Case Study” section.

Error Correction Code	Rate (Redundancy)	Burst		Single Erasure	Two Erasures
		B_{\max}	T_{worst}	T_{worst}	T_{worst}
(6, 3) RS	1/2 (100%)	3	5	3	4
RLC code	1/2 (100%)	3	5	1	3
Shifted repetition (MS)	1/2 (100%)	5	5	5	∞
Shifted RLC (ERLC)	1/2 (100%)	4	4	4	5
Concatenated code (MIDAS)	4/9 (125%)	5	5	4	5
Dual-delay code	1/2 (100%)	4	5	1	5

We assume a maximum recovery delay of $T = 5$ for the burst-erasure channel and compute the maximum correctable burst length. We also indicate the achievable delays over the isolated-erasure channel, with $N = 1$ and $N = 2$ erasures. The codes indicated in parentheses are generalizations of the codes discussed in this section.

both correct from $N = 2$ isolated erasures, require only a slightly larger delay. Similarly, if we look at a recovery delay of $T = 11$, we see that the streaming codes—shifted-repetition (MS), shifted-RLC (ERLC), and concatenated (MIDAS) codes—can recover from bursts of lengths 11, 10, and 9, respectively, compared to a $B = 6$ for the RLC code of the same redundancy.

We conclude by noting that while this section considers very simple channel models throughout the discussion, the insights gleaned from this study are valuable over more realistic channel models. This will be validated in the “Simulation Results” section, where we show how the streaming codes can outperform conventional codes over Gilbert–Elliott channels as well as real packet traces.

General code constructions

Here, we extend the streaming code constructions discussed in the previous section to general parameters. We first examine MS codes. Recall that these codes are a generalization of shifted-repetition codes. MS codes achieve optimal error correction over the burst-erasure channel. However, they cannot recover from even $N = 2$ isolated erasures. We then outline two approaches—the ERLC code and the MIDAS code—that are also robust to isolated losses.

MS codes

MS codes were introduced in [24] and [25] and shown to achieve maximum burst-correction capability for a given rate and delay. The original constructions of MS codes in [24] and [25] were based upon interleaved block codes. A modification was suggested in [13] that did not use the block code construction. Instead, the MS code was constructed using an RLC code and a repetition code as its constituent codes. We follow this approach, as it is simpler to describe and generalizes naturally to the robust extensions.

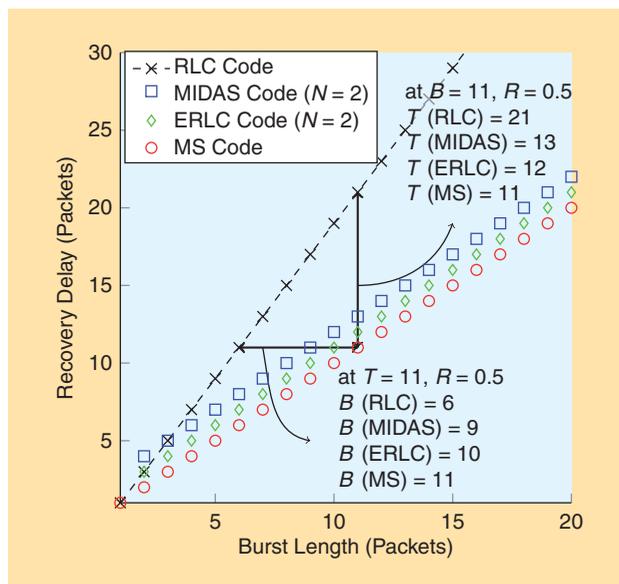


FIGURE 8. The achievable recovery delays for different burst lengths using FEC at rate 1/2. Sequential recovery codes (such as MS codes) incur a much lower delay when compared to simultaneous recovery codes (such as RS codes and RLC codes) for a given burst length.

Before explaining the detailed construction, we make the following remark about its constituent codes. The shifted-repetition code is an intrapacket code, so called because it does not combine symbols belonging to different source packets. It sequentially recovers the source packets, but its rate is fixed at 1/2. By contrast, the RLC code is an interpacket code, as it combines symbols across different source packets [see (4)]. This construction allows for a flexible rate-delay tradeoff but achieves only simultaneous recovery. In the MS code construction, we combine the contributions of both the RLC code and the repetition code, as illustrated in Figure 9.

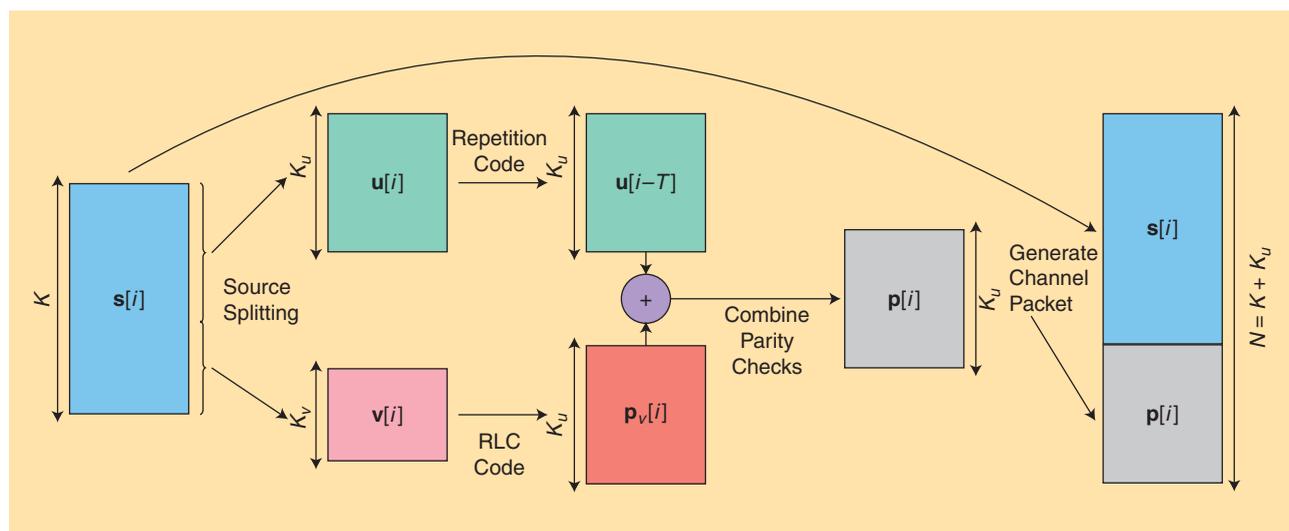


FIGURE 9. A block diagram illustrating the encoding steps of an MS code. The source packet is first split into two subpackets, and a different code is applied to each subpacket. The resulting parity checks are then combined to form the overall parity-check packet. Finally, the parity-check packet is appended to the source packet to generate the channel packet.

Encoder

- 1) *Source splitting*: Split each source packet into two sub-packets $\mathbf{u}[i]$ and $\mathbf{v}[i]$ of sizes K_u and K_v , respectively, where $K_u + K_v = K$, i.e.,

$$\mathbf{s}[i] = (\mathbf{u}[i], \mathbf{v}[i]).$$

- 2) *RLC code*: Apply a rate $K_v/(K_v + K_u)$ RLC code to the $\mathbf{v}[\cdot]$ stream of subpackets to generate parity-check packets $\mathbf{p}_v[\cdot]$ of size K_u .
- 3) *Repetition code*: Apply a shifted-repetition code to the $\mathbf{u}[\cdot]$ subpackets.
- 4) *Parity combination*: Combine the $\mathbf{p}_v[\cdot]$ parity-check packets with the repeated $\mathbf{u}[\cdot]$ subpackets after shifting the latter by T time slots to generate the overall parity-check packets,

$$\mathbf{p}[i] = \mathbf{p}_v[i] + \mathbf{u}[i - T].$$

- 5) *Channel packet*: Generate the channel packets by appending the overall parity checks to the source packets, i.e., $\mathbf{x}[i] = (\mathbf{s}[i], \mathbf{p}[i])$ is the packet transmitted at time i and is of size $N = K + K_u$.

Rate analysis

In the above construction, we may select any value of K_u and K_v such that their ratio is $K_u/K_v = B/(T - B)$. The overall rate is given by $R = (K_u + K_v)/(2K_u + K_v) = T/(T + B)$. We next explain how the code can recover from a burst of length B with a delay of T .

Decoder

Consider a channel that introduces an erasure burst of length B in the interval $[0, B - 1]$, as shown in Figure 10. The decoder proceeds in two steps.

- 1) *Simultaneous recovery*: The decoder subtracts the unerased $\mathbf{u}[B - T], \dots, \mathbf{u}[-1]$ subpackets from the corresponding parities $\mathbf{p}[B], \dots, \mathbf{p}[T - 1]$ to recover the parity-check packets $\mathbf{p}_v[B], \dots, \mathbf{p}_v[T - 1]$. These $T - B$ parities, each consisting of K_u symbols, suffice to recover the B erased $\mathbf{v}[\cdot]$ symbols, since $B \cdot K_v = (T - B) \cdot K_u$ holds.

- 2) *Sequential recovery*: Upon recovering $\mathbf{v}[0], \dots, \mathbf{v}[B - 1]$ at time $T - 1$, the decoder can compute $\mathbf{p}_v[T]$, subtract it from $\mathbf{p}[T] = \mathbf{p}_v[T] + \mathbf{u}[0]$, and in turn recover $\mathbf{u}[0]$ at time T . Similarly, the decoder can use $\mathbf{p}[T + 1], \dots, \mathbf{p}[T + B - 1]$ to sequentially recover $\mathbf{u}[1], \dots, \mathbf{u}[B - 1]$ with a delay of T packets.

Hence, $\mathbf{s}[i] = (\mathbf{u}[i], \mathbf{v}[i])$ for $i \in \{0, \dots, B - 1\}$ are recovered at time $i + T$.

We summarize the error-correction property of the MS code in the following section [24], [25].

Theorem 2 (error-correction properties of MS codes at a given maximum delay)

Given a rate R and delay T , the MS code can recover from a burst-erasure channel of maximum length B or an isolated-erasure channel with N erasures provided that

$$B \leq \min\left(1, \frac{1-R}{R}\right)T, \tag{8}$$

$$N \leq 1. \tag{9}$$

Furthermore, the upper bound on B in (8) is the maximum value that can be attained by any code of rate R and delay T .

Robust extensions of MS codes

As shown in Figure 10, the MS code splits the source packet into two groups, i.e., $\mathbf{s}[i] = (\mathbf{u}[i], \mathbf{v}[i])$. It applies a shifted-repetition code to $\mathbf{u}[i]$ and an RLC code to $\mathbf{v}[i]$ to generate the parity-check packet $\mathbf{p}_v[i] + \mathbf{u}[i - T]$. The main weakness of this construction is the shifted-repetition code applied to the $\mathbf{u}[\cdot]$ packets. When there are two isolated losses, at time $t = 0$ and $t = T$, the MS code fails to recover the subpacket $\mathbf{u}[0]$. We discuss two ways in which these codes can be made robust to correct from isolated losses.

MIDAS codes

The main idea in the MIDAS construction is to apply an additional RLC code of rate $K_u/(K_u + K_r)$ to the $\mathbf{u}[i]$ subpackets. This generates a new set of parity-check packets $\mathbf{p}_u[i]$ consisting of K_r symbols. These are then

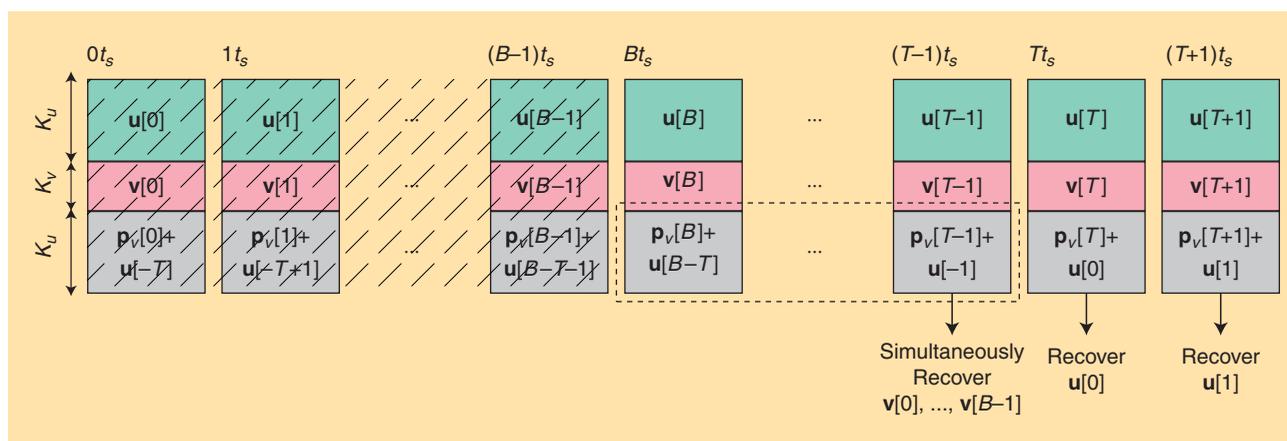


FIGURE 10. An illustration of the decoding steps in an MS code. Each column denotes a channel packet transmitted at the time index shown above it.

appended to the MS code. Thus, the transmitted channel packet is of the form $\mathbf{x}[i] = (\mathbf{u}[i], \mathbf{v}[i], \mathbf{u}[i - T] + \mathbf{p}_v[i], \mathbf{p}_u[i])$. By judiciously selecting K_r , one can achieve any $N \leq B$ (see [13]). This construction is a generalization of the concatenated code discussed in the previous section. The following result from [13] characterizes the performance of these codes.

Theorem 3 (error-correction properties of MIDAS codes at a given maximum delay)

Given a rate R and delay T , there exists a MIDAS code that can recover from a burst-erasure channel of maximum length B or an isolated-erasure channel with N erasures, provided that B and N , with $1 \leq N \leq B$, satisfy the following:

$$\left(\frac{R}{1-R}\right)B + N \leq T. \quad (10)$$

Unlike the case of RLC codes in Theorem 1 where $N = B$ and the case of MS codes in Theorem 2 where $N = 1$, the family of MIDAS codes can achieve any value of $N \in [1, B]$ in Theorem 3. Equation (10) governs the tradeoff between the burst-error and isolated-error correction capabilities of MIDAS codes for a given rate and delay. As the value of N increases, the value of B must decrease and vice versa. Finally, it is also established [13] that the tradeoff in (10) is within one unit of the optimal delay.

ERLC codes

In this approach, we replace the repetition code in the MS code with an RLC code. As with the MS codes, we split the source packet $\mathbf{s}[i] = (\mathbf{u}[i], \mathbf{v}[i])$ of size K_u and K_v symbols, respectively. We apply an RLC code to the $\mathbf{v}[i]$ packets as before to generate parity checks $\mathbf{p}_v[i]$ consisting of K_u symbols. However, we substitute the rate-1/2 repetition code applied to the $\mathbf{u}[i]$ packets with a rate-1/2 RLC code to generate parity-check packets $\mathbf{p}_u[i]$ consisting of K_u symbols. The channel packet transmitted at time i is expressed as $\mathbf{x}[i] = (\mathbf{u}[i], \mathbf{v}[i], \mathbf{p}_v[i] + \mathbf{p}_u[i - \Delta])$, where $\Delta \in [0, T]$ denotes the shift applied to the $\mathbf{p}_u[\cdot]$ stream. By judiciously selecting the value of Δ , we can trade off the burst-error correction and the isolated-error correction capability of this code [15].

Theorem 4 (error-correction properties of ERLC at a given maximum delay)

Consider an ERLC code of rate R , delay T , and shift Δ that satisfies $\Delta \geq R(T + 1)$. For $R \geq 1/2$, the ERLC code can recover from a burst-erasure channel with maximum burst length B or from an isolated-erasure channel with a maximum of N erasures, provided that

$$B \leq \frac{1-R}{R} \Delta, \quad (11)$$

$$N \leq \frac{1-R}{R} (T - \Delta) + 1. \quad (12)$$

In the broader literature, there has been a long-standing interest in packet-level convolutional codes for burst-error correction.

Remark 3

In the ERLC construction, the choice of the shift Δ is a design parameter. By varying the value of Δ , we can attain a tradeoff between the burst-error and isolated-error correction capabilities of the code [15].

From Theorems 3 and 4, at rate $R = 1/2$, ERLC codes achieve larger values of B and

N than MIDAS at a given T . ERLC codes are also shown to outperform MIDAS codes on patterns consisting of a burst followed by isolated losses (compare [14]). We will see that these advantages of ERLC codes also lead to improved performance in simulations over the statistical channel models. But before presenting the simulation results, we provide a survey of the existing literature on streaming codes.

Literature survey

Having discussed some of the basic streaming code constructions in the previous sections, we provide a survey of the existing literature in this area. In the broader literature, there has been a long-standing interest in packet-level convolutional codes for burst-error correction (see [21], [26]–[30] and the references therein). However, these references do not impose the decoding delay constraint and focus only on error recovery. The streaming setup in Figure 2 was introduced, to our knowledge, by Martinian and Sundberg [24]. The class of MS codes for the burst-erasure channel that we discussed in the “MS Codes” section was also presented in [24]. These were further developed in [20] and [25], where explicit code constructions were provided for all feasible burst lengths and decoding delays. The constructions in these works were based on a two-stage approach. A low-delay block code was first constructed and then interleaved to construct a convolutional code. Later, [13] provided an alternative approach that did not require the block code construction but directly constructed the convolutional code using an RLC code and a repetition code as constituent codes. This approach was outlined in the “MS Codes” section.

While MS codes achieve the optimal burst erasure correction capability, they are sensitive to other loss patterns. In [13]–[15], a sliding window channel model with burst and isolated erasures is introduced, and the MIDAS and ERLC codes are introduced in these works. A fundamental tradeoff between the burst-erasure and isolated-erasure correction properties of any code is established. This framework is used to establish certain optimality properties of the proposed codes. Our discussion of MIDAS and ERLC codes in the “Robust Extensions of MS Codes” section is based on these references.

Throughout this tutorial article, we restrict our attention to the case in which one source packet arrives in each time slot and one channel packet must be transmitted in each slot. References [13] and [31]–[33] consider the case where the source arrival and channel transmission rates are mismatched. In particular, $M > 1$ channel packets must be transmitted by the encoder between two successive source packets. References [13] and [31] consider the decoding delay in terms of the source packets and characterize the capacity for the case of burst-erasure

channels. The associated code constructions are also based on a layering scheme, as in the MS constructions. The optimality of these codes is established for the burst-erasure channel model. References [32] and [33] study a similar setup, in which the decoding delay is with respect to channel packets. For the burst erasure model, diagonally interleaved block codes are shown to be optimal when gaps between successive bursts are sufficiently small. For the independent and identically distributed (i.i.d.) erasure model, a family of time-invariant intrasession codes is proposed, with a performance that is close to an upper bound.

References [34] and [35] consider a model where the transmitter and receiver are connected through multiple parallel links. Each link is assumed to be a burst-erasure channel that introduces a burst of maximum length B . The capacity is characterized in some special cases, and joint coding across the subchannels is required to attain the capacity. Reference [36] considers the setup when the channel between the source and destination is modeled using a linear transfer matrix and is subjected to rank losses. Convolutional coding analogs of rank error correcting codes are proposed that maximize a new distance metric known as the *maximum column sum rank*. In reference [37], the problem of having multiple erasure bursts within each coding block is studied. It is shown that the recovery delay depends not only on the number of bursts within a coding block but on whether the source symbols are encoded causally or noncausally.

In other related works, [38]–[41] study a multicast extension of [24] and [25] involving two users and a common source stream. The stronger receiver's channel introduces shorter bursts, and in turn the decoding delay is required to be smaller. The weaker receiver's channel introduces longer bursts, and the decoding delay can be longer. Such codes can also be used in applications where the decoding delay can vary based on channel conditions. The construction of these codes involves embedding the parity checks of two single-user MS codes in a careful manner to simultaneously satisfy the constraints of both receivers.

In the broader literature, unequal error protection for multimedia streaming has been widely studied (see, e.g., [42]–[44] and the references therein). In [45], the authors proposed a new scheme in streaming models with feedback, which combines the benefits of network coding and ARQ by acknowledging degrees of freedom instead of original packets. In [46]–[48], real-time streaming over blockage channels with delayed feedback is studied. A multiburst transmission protocol is proposed that achieves a nontrivial tradeoff between the delay and throughput within this framework.

Simulation results

In this section, we study the performance of different FEC codes over Gilbert–Elliott channels as well as real packet traces. In our simulations, we fix the rate of the code to be $R = 1/2$. In practice, error correction may be invoked only on a subset of packets. For example, a large fraction of packets in an audio stream correspond to silence periods. These packets clearly do not need error control. Second, error control may

be only adaptively invoked when the channel conditions require it [11]. Such approaches can substantially reduce the overhead from FEC packets. The maximum recovery delay used in this section is $T = 5$ and $T = 12$, as suggested in the “Impact on Applications” section. Furthermore, the packet loss rates (PLRs) we consider are in the interval 10^{-3} to 10^{-6} . The former loss rate will result in a playback disruption every few seconds; the latter loss rate will produce a playback disruption only once every half hour or so.

Gilbert–Elliott channel

A Gilbert–Elliott channel is a two-state Markov model. In the good state each channel packet is lost with probability ϵ , whereas in the bad state each channel packet is lost with probability 1. The average loss rate of the Gilbert–Elliott channel is given by

$$\Pr(\alpha, \beta, \epsilon) = \frac{\beta}{\beta + \alpha} \epsilon + \frac{\alpha}{\alpha + \beta}, \quad (13)$$

where α and β denote the transition probabilities from the good state to the bad state and vice versa. As long as the channel stays in the bad state, the channel behaves as a burst-erasure channel. The length of each burst is a geometric random variable of mean $1/\beta$. When the channel is in the good state, it behaves as an i.i.d. erasure channel with an erasure probability of ϵ . The gap between two successive bursts is also a geometric random variable with a mean of $1/\alpha$. Finally, $\epsilon = 0$ results in a Gilbert–Elliott channel [49], which results in burst losses only.

In Figure 11, we fix $\alpha = 5 \times 10^{-4}$ and vary both β and ϵ of the Gilbert–Elliott channel to achieve different mean burst lengths (on the y axis) and i.i.d. loss rates (on the x axis). Each point corresponds to a single realization of 10^8 packets. We use rate $R = 1/2$ shifted-repetition, shifted-RLC, and RLC codes from the “Numerical Comparisons” section and set the maximum delay to $T = 5$ packets.

The code with the minimum residual loss rate at a given mean burst length $1/\beta$ and i.i.d. loss percentage 100ϵ is marked in Figure 11. It turns out that there are three main regions and that each code dominates in one. As expected, MS codes outperform other codes when the mean burst lengths are high compared to i.i.d. loss rates between bursts. In the other extreme, RLC codes are the best. Finally, ERLC codes, which can recover bursts slightly longer than RLC codes can and more i.i.d. losses than MS codes can dominate in a region between the two extremes. This shows that an application can gainfully switch between the three codes, depending on the expected channel characteristics.

Figure 12 illustrates the results of another experiment over Gilbert–Elliott channels with parameters given in the second column of Table 2. We ran simulations over 31 realizations of a Gilbert–Elliott channel, each of length 10^8 packets. We set $\alpha = 5 \times 10^{-4}$ and $\beta = 0.4$ in all realizations and varied $\epsilon \in [0, 3] \times 10^{-2}$ across realizations. In Figure 12, the channel loss rate, $\Pr(\alpha, \beta, \epsilon)$, is plotted on the x axis, whereas the residual loss probability of different streaming codes is plotted

on the y axis. We assume that the rate of all codes is $R = 1/2$ and the delay is $T = 12$. The achievable values of N and B over the isolated-erasure channel and the burst-erasure channels are shown in Table 2.

The curve indicated by the black line and marked with \times signs corresponds to the RLC codes. These codes achieve the largest value of N among all the codes in Table 2. This explains the relatively constant performance as ϵ is increased. The bottleneck for these codes is long erasure bursts. In particular, in Table 2 these codes achieve a much smaller value of B and hence incur significant packet losses due to long bursts. The curve indicated by the light-blue line and marked with $+$ signs corresponds to RS block codes. These codes achieve the same value of N and B and hence exhibit a pattern similar to RLC codes. However, as discussed before, they are not adaptive and are weaker to nonideal erasure patterns.

The red plot marked with circles corresponds to MS codes. These codes are optimal for the burst-erasure channel and achieve the largest value of B among all the codes in Table 2. However, they achieve only $N = 1$, and hence their performance is very sensitive to isolated erasures in the good state. In particular, as ϵ increases, the performance deteriorates quickly.

The dark-blue plot marked with squares corresponds to the MIDAS codes. These codes can balance between the values of B and N and are able to correct both isolated erasures in the good state and longer burst losses in the bad state. MIDAS codes combine the advantages of MS and RLC codes.

The green plot marked with diamonds is for the ERLC codes. Similar to MIDAS codes, ERLC codes can balance between the values of B and N . The improvement in loss rate achieved by ERLC codes is due to their capability to partially recover from some nonideal patterns consisting of burst and isolated erasures in the same decoding window (compare [14]). Overall, Figure 12 demonstrates the improvements that

different streaming codes can realize over traditional RS and RLC codes.

Figure 13 studies the effect of increasing delay on different codes. We consider a simulation over the Gilbert–Elliott channel, with $\alpha = 5 \times 10^{-4}$, $\beta = 0.4$, and $\epsilon = 4 \times 10^{-3}$. We plot the residual loss rate of different codes versus the allowed delay T in the range of five to 25 packets. At each delay T and rate $R = 1/2$, RLC and MS codes can achieve only a single pair of (B, N) values, whereas ERLC and MIDAS codes can achieve a set of pairs. The selected pairs, shown in Table 3, correspond to the minimum residual loss rate among all pairs. As the allowed delay increases, the isolated erasures and/or burst-erasure correction capabilities can be enhanced, as shown in Table 3.

Table 2. The channel and code parameters used in the simulations.

	Figure 11		Figure 12	
α	5×10^{-4}		5×10^{-4}	
β	[1/3, 1)		0.4	
ϵ	$[0, 3 \times 10^{-2}]$		$[0, 3 \times 10^{-2}]$	
Channel length	N/A		5×10^7	
Rate R	1/2		1/2	
Delay T	5		12	
	N	B	N	B
RS	–	–	6	6
RLC	3	3	6	6
MS	1	5	1	12
MIDAS	–	–	2	10
ERLC	2	4	2	11

The values of B and N indicate, respectively, the maximum burst length and the number of isolated losses that can be corrected by each code.

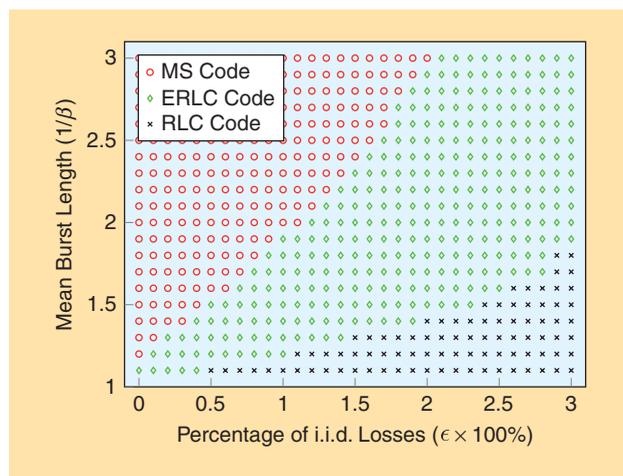


FIGURE 11. A numerical comparison over a Gilbert–Elliott channel at $\alpha = 5 \times 10^{-4}$. We vary $\beta \in [1/3, 1)$ to achieve mean burst lengths of $1/\beta$ on the y axis and $\epsilon \in [0, 3 \times 10^{-2}]$ to achieve i.i.d. loss percentages of $100\epsilon\%$ on the x axis. At each point, we indicate the code that achieves the minimum residual PLR.

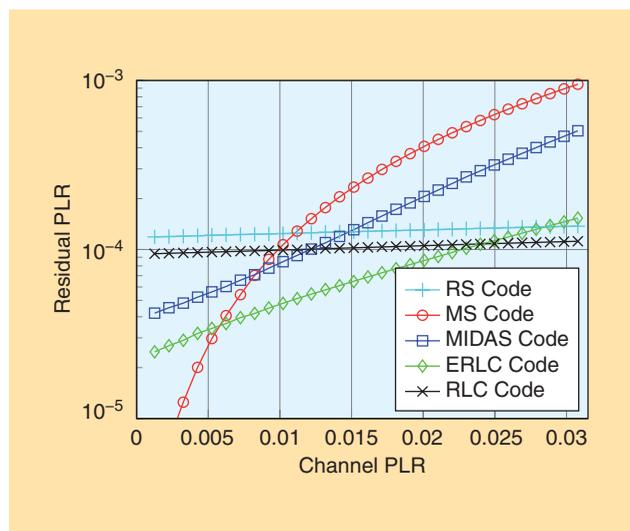


FIGURE 12. A simulation over a Gilbert–Elliott channel model, with $(\alpha, \beta, \epsilon) = (5 \times 10^{-4}, 0.4, [0, 0.03])$. The rate for all codes is $R = 1/2$ and the delay is $T = 12$ packets.

However, only B of the MS code can be increased and not N . Starting at $T = 11$, which corresponds to $B = 11$ for the MS code, its residual loss rate is dominated by the isolated erasure patterns, and hence its residual loss rate saturates. On the other hand, as the allowed delay increases, selecting the right values of N and B (compare Table 3) for ERLC and MIDAS codes helps to improve their performance over RLC and MS codes, which can achieve only a single (B, N) pair for a given T .

Real packet traces

In this section, we validate our results over real packet traces. We simulate the RLC, shifted-repetition, and shifted-RLC codes over the data set in [50]. This data set consists of over 150 million packets collected over a wireless sensor network while varying multiple parameters, for example, packet interarrival time, payload size, and distance between nodes. We consider packets with an interarrival time equal to 20 ms, since this models most VoIP applications. There are a total of 18.75 million packets, with a loss rate of 8.3%. We use the same codes shown in Figure 12, whose parameters are indicated in the first column of Table 2, and set the delay to $T = 5$ packets or 100 ms. We divide the traces into nonoverlapping

windows of length 15,000 packets each, i.e., 5 min of audio. The window length of 15,000 packets is chosen to be the approximate coherence time of the channel. Out of the 1,250 windows, 133 are loss free and 139 contain long bursts (> 50 packets). We focus on the remaining 978 windows with moderate mean burst lengths, because we believe the long bursts are due to outages and/or link failures and no FEC can recover from such patterns.

Figure 14 indicates the code with the minimum residual PLR for each of the 978 considered windows. We plot the average nonbursty PLR in each window (the sum of isolated losses divided by the length of the window) on the x axis versus the average burst length in each window (considering any two or more consecutive erasures as a burst) on the y axis. Interestingly, each of the three simulated codes dominates in a different region.

- 1) Windows with a short mean burst length (less than 2.5), i.e., close to the x axis in Figure 14. In these windows, the isolated losses are the dominant erasure patterns. RLC codes are designed for such channels and achieve the minimum loss rate among all the simulated codes.
- 2) Windows with a small number of isolated packet losses but relatively long mean burst lengths, i.e., the top left corner in Figure 14. Most of the erasures in these windows are due to bursts. Hence, the shifted-repetition (MS) code, which has the longest burst-erasure correction capability $B = 5$, achieves the minimum loss rate in the majority of such windows.
- 3) Windows that introduce a relatively balanced mixture of isolated losses and long bursts. The shifted-RLC (ERLC) code, which can recover from a longer burst $B = 4$ compared to the RLC code and from more isolated losses $N = 2$ compared to the shifted-repetition (MS) code, achieves the minimum loss rate in most of these windows, as shown in Figure 14.

Table 4 includes further results of our experiments with these traces. Each row corresponds to a subset of windows where a group of codes achieves the minimum residual loss rate. For each subset, we indicate the following:

- 1) the number of windows in the subset
- 2) the average PLR in these windows
- 3) the average nonbursty PLR corresponding to isolated losses
- 4) the average burst length (considering only $B \geq 2$)
- 5) the average of the maximum burst length across windows of the set
- 6) the average residual loss rates for all three codes—RLC, shifted repetition, and shifted RLC—in each subset.

The first three rows in Table 4 correspond to the points in Figure 14.

The values of the nonbursty PLR, mean burst length, and maximum burst length for different subsets in Table 4 confirm the results in Figure 14. ERLC code achieves the minimum loss rate in more than 60% of the windows considered. It also achieves the minimum average residual loss rate among all 978 windows. However, no single code achieves the best performance for all windows, and selecting the right code

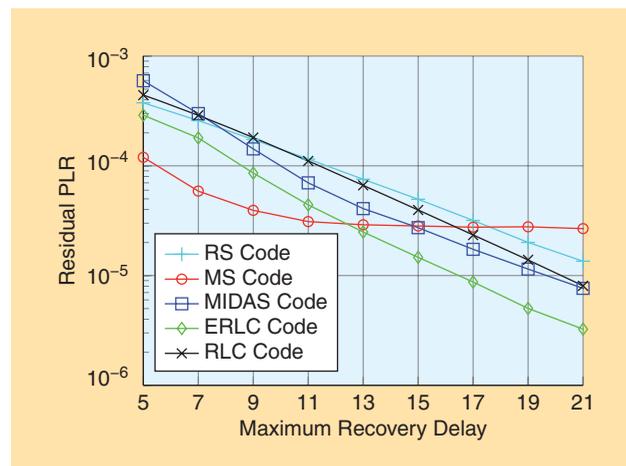


FIGURE 13. The simulation experiments for a Gilbert-Elliott channel model, with $(\alpha, \beta, \epsilon) = (5 \times 10^{-4}, 0.4, 4 \times 10^{-3})$.

Table 3. The achievable B and N for different $R = 12$ codes, as a function of T , in Figure 13.

Code	T	5	7	9	11	13	15	17	19	21
RLC	B	3	4	5	6	7	8	9	10	11
	N	3	4	5	6	7	8	9	10	11
MS	B	5	7	9	11	13	15	17	21	23
	N	1	1	1	1	1	1	1	1	1
MIDAS	B	3	5	7	9	11	13	13	14	15
	N	2	2	2	2	2	2	4	5	6
ERLC	B	4	6	8	10	12	13	14	14	15
	N	2	2	2	2	2	3	4	6	7

significantly reduces the residual loss rate in most of the cases. This suggests designing a system that adaptively selects the right code depending on the loss characteristics. According to the considered trace, one can simplify the system by alternating between only two codes. The first is ERLC, which is the best in 603 out of 978 windows (> 60%). It also achieves a loss rate that is close to that of the best code outside these windows. The second is the MS code, which uniquely achieves the minimum residual loss rate in 253 windows. Most of them lie in the top left corner of Figure 14, where long bursts are the dominant loss pattern.

Another adaptive approach can use the fact that the ERLC code is a generalization of both MS and RLC codes (compare with the “Remark 3” section). Depending on the loss statistics, mean burst length, and average loss rate, one can select the right value of the shift Δ for the chosen ERLC code. This

The main conclusion from simulation results over both statistical channels and real packet traces is that no single code achieves the best performance in all cases.

includes $\Delta = 0$, $\Delta = 4$, and $\Delta = 5$, which are the RLC, ERLC, and MS codes, respectively, shown in Figure 14 and Table 4. This will further simplify the system design, since a single code will be implemented.

The main conclusion from simulation results over both statistical channels and real packet traces is that no single code achieves the best performance in all cases. How-

ever, depending on the loss characteristics, we can estimate which code yields the best performance. Hence, we believe that by tracking the end-to-end delay conditions and the loss characteristics during a session, the system can dynamically select the right FEC code and its parameters, such as code rate, recovery delay, and burst- and isolated-correction capabilities. The system can in some cases infer the type of bottleneck (e.g., cable modems with drop tail queuing protocol frequently lead to burst losses), which can help make more informed choices about what type of losses to expect in the future and what type of FEC code would be best. Ideally, the FEC code and rate would be adapted dynamically throughout a communication session, similar to how the transmitted bit rate is dynamically varied during a session in many video applications today.

Conclusions

Interactive streaming applications require communication systems that achieve low latency and high reliability in the delivery of source packets. FEC codes provide a natural solution to these applications. However, traditional FEC codes are not designed to satisfy the low-delay and real-time requirements of these applications. As a result, many off-the-shelf codes can result in suboptimal error correction [11].

An exciting opportunity exists to develop new classes of streaming codes for interactive streaming application. This tutorial provides a survey of the current state-of-the-art constructions of such codes, uses simple illustrative examples to

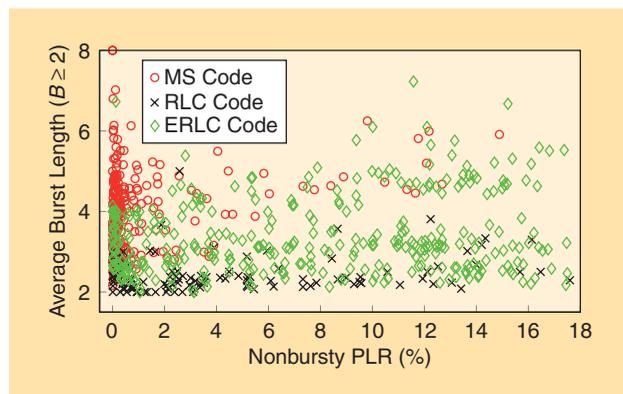


FIGURE 14. The simulation results over real packet traces. Each point in the figure corresponds to a window of length 15,000 packets, with the nonbursty PLR on the x axis and the mean burst length on the y axis. The code that achieves the minimum residual loss rate at each window is indicated with its corresponding mark. These correspond to the first three rows in Table 4.

Table 4. An analysis of the simulations over real packet traces.

Code with Minimum PLR	Indicator	Number of Windows	PLR (%)	Nonbursty PLR (%)	Mean Burst Length	Maximum Burst Length	Residual PLR (%)		
							RLC	MS	ERLC
RLC	0001	346	2.43	2.04	2.03	3.14	0.08	0.51	0.10
MS	0100	441	3.72	0.68	3.23	8.15	2.17	1.43	1.71
ERLC	0010	603	5.89	4.49	2.82	5.79	0.91	1.64	0.73
RLC and MS	0101	146	0.09	0.06	1.65	1.82	0.00	0.00	0.00
MS and ERLC	0110	184	0.12	0.05	2.09	2.27	0.02	0.00	0.00
RLC and ERLC	0011	224	0.98	0.88	1.79	2.16	0.02	0.20	0.02
All codes	0111	142	0.09	0.05	1.62	1.71	0.00	0.00	0.00
Total	1000	978	5.92	3.58	3.10	7.41	1.56	1.79	1.25

Each row corresponds to a subset of windows in which a code achieves the minimum residual loss rate. The average of the following values across such a subset of windows is also indicated: the PLR, isolated-only PLR, mean burst length, maximum burst length, and residual PLR of each code for such windows.

provide insights into these constructions, and summarizes the layered design underlying these codes.

We first explain why traditional FEC, such as RS codes, rateless codes, and RLC codes, are not ideal in streaming applications. These codes do not explicitly consider the different deadlines of different source packets. The resulting code forces the decoder to recover the erased packets simultaneously, without taking into account the different decoding deadlines.

We then discuss the class of MS codes that achieve optimal error correction over burst-erasure channels by recovering the erased source packets in a sequential fashion. The decoder is thus capable of recovering older packets with earlier deadlines before the newer packets. These codes correct burst lengths that can be twice as long as traditional codes or equivalently reduce the recovery delay by up to a factor of two for a given burst length. We then discuss two additional codes—the MIDAS codes and the ERLC codes—that sacrifice a small amount of burst-error correction capability to achieve significant improvements in robustness over the isolated-erasure channel model. We provide both specific examples and outline general constructions for these codes. We compare their performance over a variety of packet-loss sequences and also demonstrate that they achieve significant gains in simulations over statistical channel models and real packet traces.

Many promising future directions are possible. One direction is to design systems that can opportunistically select among different FEC codes, depending on the application constraints and current channel statistics, such as end-to-end delay and loss characteristics. Moreover, designing FEC codes whose recovery delay adapts to the state of the channel can be beneficial in applications using adaptive playback techniques. Also, content-aware FEC that adapts to the importance of the source stream can provide improved perceptual quality. Another valuable direction is designing low-delay FEC in the case of multiple streams with different delay constraints.

We believe this is a highly promising area for improvement in interactive voice and video communications, augmented and virtual reality applications, and various IoT use cases and hope that we have conveyed to the reader our excitement about these new opportunities.

Acknowledgments

This work was supported by a National Science and Engineering Research Council of Canada discovery grant, a Hewlett-Packard Innovation Research Program Award, a Cisco Systems grant, and an Ontario Early Researcher Award.

Authors

Ahmed Badr (abadr@ece.utoronto.ca) received his B.S. degree from Cairo University, Egypt, in 2007; his M.S. degree from Nile University, Egypt, in 2009; and his Ph.D.

By tracking the end-to-end delay conditions and the loss characteristics during a session, the system can dynamically select the right FEC code and its parameters.

degree from the University of Toronto, Canada, in 2014, all in electrical and computer engineering. From September 2007 to August 2009, he was a research assistant with the Wireless Intelligent Networks Center, Nile University. In September 2009, he was a research assistant with the Signals Multimedia and Security Laboratory at the University of

Toronto. In 2014, he assumed his current position as a post-doctoral fellow. His research interests include information theory, coding theory, and real-time communication. He is a Member of the IEEE.

Ashish Khisti (akhisti@ece.utoronto.ca) received his B.A.S. degree in engineering sciences (electrical option) from the University of Toronto, Canada, and his S.M. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT). From 2009 to 2015, he was an assistant professor in the Electrical and Computer Engineering Department at the University of Toronto. He is presently an associate professor and holds a Canada Research Chair in the same department. He is a recipient of an Ontario Early Researcher Award, the Hewlett-Packard Innovation Research Program Award, and the Harold H. Hazen Teaching Assistant Award from MIT. He is an associate editor of *IEEE Transactions on Information Theory* and was a guest editor of *Proceedings of the IEEE* (special issue on secure communications via physical-layer and information-theoretic techniques). He is a Member of the IEEE.

Wai-Tian Tan (dtan2@cisco.com) received his B.S. degree from Brown University, Providence, Rhode Island, his M.S.E.E. degree from Stanford University, California, and his Ph.D. degree from the University of California, Berkeley, all in electrical engineering. He was a researcher at Hewlett-Packard Laboratories from 2000 to 2013, working on various aspects of multimedia communications and systems. Since 2013, he has been with Cisco Systems, where he is a principal engineer in the Chief Technology and Architecture Office.

John Apostolopoulos (johnapos@cisco.com) received his B.S., M.S., and Ph.D. degrees from the Massachusetts Institute of Technology (MIT). He is vice president and chief technology officer of Cisco's Enterprise Segment and leads Cisco's Innovation Labs, covering the Internet of Things, wireless, application-aware networking, multimedia networking, indoor location-based services, and deep learning for visual analytics. Previously, he was lab director for the Mobile and Immersive Experience Lab at Hewlett-Packard Laboratories. His work has spanned novel mobile devices and sensing, client/cloud multimedia computing, multimedia networking, immersive video conferencing, software-defined networks, and mobile streaming media content delivery networks for all-IP (fourth generation) wireless networks. He was a consulting associate professor at Stanford University (2000–2009) and frequently lectures at MIT. He is a Fellow of the IEEE.

References

- [1] Cisco. (2014, June 10). Cisco Visual Networking Index: Forecast and methodology, 2013–2018. White Paper. [Online]. Available: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html>
- [2] W. Jiang and H. Schulzrinne, "Comparison and optimization of packet loss repair methods on VoIP perceived quality under bursty loss," in *Proc. 12th ACM Int. Workshop Network and Operating Systems Support Digital Audio and Video*, 2002, pp. 73–81.
- [3] Y. J. Liang, J. G. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: Effect of burst losses and correlation between error frames," *IEEE Trans. Circuits Systems Video Technology*, vol. 18, no. 7, pp. 861–874, Jul. 2008.
- [4] J. S. Marcus, *Designing Wide Area Networks and Internetworks: A Practical Guide*. Reading, MA: Addison-Wesley, 1999.
- [5] *One-Way Transmission Time*, International Telecommunication Union Recommendation G.114, 2000.
- [6] M. Suznjevic and J. Saldana. (2015, Dec.). Delay limits for real-time services. [Online]. Available: <https://tools.ietf.org/id/draft-suznjevic-dispatch-delay-limits-00.html>
- [7] T. Stockhammer and M. Hannuksela, "H.264/AVC video for wireless transmission," *Wireless Commun.*, vol. 12, pp. 6–13, Aug. 2005.
- [8] National Science Foundation. (2015, Mar. 26–27). NSF workshop on achieving ultra-low latencies in wireless networks. [Online]. Available: <http://inlab.lab.asu.edu/nsf>
- [9] M. Luby, "LT codes," in *Proc. Annu. IEEE Symp. Foundations Computer Science*, 2002, pp. 271–282.
- [10] A. Shokrollahi, "Raptor codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2551–2567, 2006.
- [11] T. Huang, P. Huang, K. Chen, and P. Wang, "Could Skype be more satisfying? A QoE-centric study of the FEC mechanism in an Internet-scale VoIP system," *IEEE Netw.*, vol. 24, no. 2, pp. 42–48, 2010.
- [12] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *SIAM J. Appl. Math.*, vol. 8, pp. 300–304, June 1960.
- [13] A. Badr, P. Patil, A. Khisti, W. Tan, and J. Apostolopoulos, "Layered constructions for low-delay streaming codes," *IEEE Trans. Inf. Theory*, vol. 63, no. 1, pp. 111–141, 2016.
- [14] A. Badr, A. Khisti, W. Tan, and J. Apostolopoulos, "Streaming codes with partial recovery over channels with burst and isolated erasures," *IEEE J. Sel. Topics Signal Process.*, vol. 9, pp. 501–516, Apr. 2015.
- [15] A. Badr, A. Khisti, W. Tan, and J. Apostolopoulos, "Streaming codes for channels with burst and isolated erasures," in *Proc. Int. Conf. Computer Communications (INFOCOM)*, Turin, Italy, 2013, pp. 2850–2858.
- [16] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [17] Z. Li, A. Khisti, and B. Girod, "Forward error protection for low-delay packet video," in *Proc. 18th Int. Packet Video Workshop*, Dec. 2010, pp. 1–8.
- [18] P. Elias, "Coding for noisy channels," *IRE Convention Rec.*, vol. 4, pp. 37–46, Mar. 1955.
- [19] G. Joshi, "On playback delay in streaming communication," M.S. thesis, Massachusetts Institute of Technology, Cambridge, 2012.
- [20] E. Martinian, "Dynamic information and constraints in source and channel coding," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 2004.
- [21] H. Gluesing-Luerssen, J. Rosenthal, and R. Smarandache, "Strongly-MDS convolutional codes," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 584–598, 2006.
- [22] E. M. Gabidulin, "Convolutional codes over large alphabets," in *Proc. Int. Workshop Algebraic Combinatorial and Coding Theory*, Varna, Bulgaria, 1988, pp. 80–84.
- [23] J. C. Bolot, S. Fosse-Parisis, and D. Towsley, "Adaptive FEC-based error control for Internet telephony," vol. 3, pp. 1453–1460, Mar. 1999.
- [24] E. Martinian and C. W. Sundberg, "Burst erasure correction codes with low decoding delay," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2494–2502, 2004.
- [25] E. Martinian and M. Trott, "Delay-optimal burst erasure code construction," in *Proc. Int. Symp. Information Theory (ISIT)*, Nice, France, July 2007.
- [26] M. Arai, A. Yamamoto, A. Yamaguchi, S. Fukumoto, and K. Iwasaki, "Analysis of using convolutional codes to recover packet losses over burst erasure channels," in *Proc. IEEE Pacific Rim Int. Symp. Dependable Computing*, 2001, pp. 258–265.
- [27] H. Deng, M. Kuijper, and J. Evans, "Burst erasure correction capabilities of $(n, n-1)$ convolutional codes," in *Proc. IEEE Int. Conf. Communications*, 2009, pp. 1–5.
- [28] B. M. Kurkoski, P. H. Siegel, and J. K. Wolf, "Analysis of convolutional codes on the erasure channel," in *Proc. Int. Symp. Information Theory (ISIT)*, 2004, p. 460.
- [29] R. Smarandache, H. Gluesing-Luerssen, and J. Rosenthal, "Strongly MDS convolutional codes, a new class of codes with maximal decoding capability," in *Proc. Int. Symp. Information Theory (ISIT)*, Lausanne, Switzerland, 2002, p. 426.
- [30] V. Tomas, J. Rosenthal, and R. Smarandache, "Decoding of convolutional codes over the erasure channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 90–108, 2012.
- [31] P. Patil, A. Badr, and A. Khisti, "Delay-optimal streaming codes under source-channel rate mismatch," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, 2013, pp. 2094–2099.
- [32] D. Leong and T. Ho, "Erasure coding for real-time streaming," in *Proc. Int. Symp. Information Theory (ISIT)*, 2012, pp. 289–293.
- [33] D. Leong, A. Qureshi, and T. Ho, "On coding for real-time streaming under packet erasures," in *Proc. Int. Symp. Information Theory (ISIT)*, 2013, pp. 1012–1016.
- [34] D. Lui, A. Badr, and A. Khisti, "Streaming codes for a double-link burst erasure channel," in *Proc. Canadian Workshop on Information Theory (CWIT)*, 2011, pp. 147–150.
- [35] D. Lui, "Coding theorems for delay sensitive communication over burst-erasure channels," M.S. thesis, University of Toronto, Ontario, Canada, Aug. 2011.
- [36] R. Mahmood, A. Badr, and A. Khisti, "Convolutional codes with maximum column sum rank for network streaming," in *Proc. IEEE Int. Symp. Information Theory*, Hong Kong, China, June 14–19, 2015, pp. 2271–2275.
- [37] Z. Li, A. Khisti, and B. Girod, "Correcting erasure bursts with minimum decoding delay," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, 2011, pp. 33–39.
- [38] A. Khisti and J. Singh, "On multicasting with streaming burst-erasure codes," in *Proc. Int. Symp. Information Theory (ISIT)*, 2009, pp. 2887–2891.
- [39] A. Badr, A. Khisti, and E. Martinian, "Diversity Embedded Streaming Erasure Codes (DE-ScO): Constructions and optimality," *IEEE J. Sel. Areas Commun.*, vol. 29, pp. 1042–1054, May 2011.
- [40] A. Badr, D. Lui, and A. Khisti, "Multicast Streaming Codes (Mu-ScO) for burst erasure channels," in *Proc. Allerton Conf. Communication, Control, and Computing*, 2010, pp. 383–390.
- [41] A. Badr, D. Lui, and A. Khisti, "Streaming codes for multicast over burst erasure channels," *IEEE Trans. Inf. Theory*, vol. 61, pp. 4181–4208, Aug. 2015.
- [42] A. Nafaa, T. Taleb, and L. Murphy, "Forward error correction strategies for media streaming over wireless networks," *IEEE Commun. Mag.*, vol. 46, no. 1, pp. 72–79, 2008.
- [43] U. Horn, K. Stuhlmüller, M. Link, and B. Girod, "Robust Internet video transmission based on scalable coding and unequal error protection," *Signal Process.: Image Commun.*, vol. 15, no. 1, pp. 77–94, 1999.
- [44] N. Rahnavard, B. N. Vellambi, and F. Fekri, "Rateless codes with unequal error protection property," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1521–1532, 2007.
- [45] J. K. Sundararajan, D. Shah, and M. Medard, "ARQ for network coding," in *Proc. Int. Symp. Information Theory (ISIT)*, 2008, pp. 1651–1655.
- [46] H. Yao, Y. Kochman, and G. W. Wornell, "A multi-burst transmission strategy for streaming over blockage channels with long feedback delay," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 10, pp. 2033–2043, 2011.
- [47] H. Yao, Y. Kochman, and G. W. Wornell, "On delay in real-time streaming communication systems," in *Proc. Allerton Conf. Commun., Control, and Computing*, 2010, pp. 420–426.
- [48] H. Yao, Y. Kochman, G. W. Wornell, "Delay-throughput tradeoff for streaming over blockage channels with delayed feedback," in *Proc. IEEE Military Communication Conf. (MILCOM)*, 2010, pp. 1953–1958.
- [49] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, pp. 1253–1265, Sept. 1960.
- [50] S. Fu and Y. Zhang. (2015, Apr.). CRAWDAD dataset due/packet-delivery (v. 2015-04-01)—the delay traceset. [Online]. Available: <http://crawdad.org/due/packet-delivery/20150401/delay>
- [51] A. Badr, A. Khisti, W. Tan, X. Zhu, and J. Apostolopoulos, "FEC for VoIP using dual-delay streaming codes" in *Proc. IEEE Int. Conf. Computer Communications (INFOCOM)*, Atlanta, May 2017, to be published.

Konstantinos P. Michmizos, Blerta Lindqvist, Stephen Wong, Eric L. Hargreaves, Konstantinos Psychas, Georgios D. Mitsis, Shabbar F. Danish, and Konstantina S. Nikita

Computational Neuromodulation: Future Challenges for Deep Brain Stimulation

Over the past two decades, deep brain stimulation (DBS) has been leading a renaissance of neurosurgical treatments for neurological and neuropsychiatric disorders. DBS has become an established adjunct therapy for movement and mood disorders that, despite maximal medical treatment, remain sufficiently debilitating to warrant the risks of brain surgery [1]. The procedure has been approved by the U.S. Food and Drug Administration (FDA) for essential tremor (ET), Parkinson's disease (PD), dystonia, and obsessive compulsive disorder, and the growing spectrum of treatable conditions is expanding to pain and major depression, among others. Interestingly, the large phenomenological variance of the treatable symptoms that span the motor and affective domains is addressed by the same therapeutic principle: similarly to how a cardiac pacemaker works, a medical device called a *neurostimulator* sends frequent (50–250 Hz) electrical pulses to electrodes implanted into a subcortical nucleus associated with the disorder. Despite its simplicity, the procedure, when applied accurately, may alleviate symptoms of complicated diseases.

After 20 years of clinical practice and a variety of hypotheses formulated at the local or the network scale, the physiological mechanisms of DBS remain unclear. Although the surgical implantation procedure offers a

unique opportunity to record in vivo neural signals as close to their generators as possible, the recording conditions significantly vary depending on the intrinsic variability of the brain, the divergence in structural changes caused by the underlying neuropathophysiology, the compensation mechanisms that each brain has possibly developed, and the long-term administration of medication in the patients on whom were operated. Consequently, DBS

improvement has been hampered by stagnation in discovering personalized and dynamic methodologies that can leverage the intranuclear neural signals to address the highly diverse clinical phenotype and the fluctuating symptom severity. This is about to change as recently introduced DBS systems create new frontiers for the neural signal processing community. In this article, we discuss the basic principles and challenges faced by the new technological advances in DBS and describe the race toward personalizing therapy to each patient's clinical state.

Neural signals drive automatic detection of deep structures inside the human brain

The DBS implantation procedure is typically guided by microelectrode recordings (MERs) of the neural activity at different subcortical depths inside and

outside the nucleus (Figure 1). The activity is mapped via one or more microelectrodes traveling along the putative implantation path, and the resultant pattern of neural spikes is transduced to

audio. When a neurophysiologist acoustically verifies the pattern of multiunit spikes that corresponds to the entry/exit of the implantation target, the recording electrode is removed and a stimulation electrode is implanted along this trajectory.

The process gives intraoperative access to two important neurophysiological signals from deep structures: 1) the local field potential (LFP), which is the low frequency content (up to 100 Hz) of the MERs representing the synchronized oscillatory activity mainly at the dendrites of neurons up to 3 mm away from the electrode, and 2) the multiunit activity (MUA), which is the high frequency content of the MERs representing the neural spiking patterns from neurons with a distance 100–300 μm . Although our study on the LFP-MUA relationship has revealed interesting nonlinear correlations between the two signals [2], their functional interconnection remains unclear, and they are typically treated as signals carrying different types of information. Nevertheless, incorporating LFP- and MUA-derived features into neural area classifiers has supported the

Over the past two decades, deep brain stimulation has been leading a renaissance of neurosurgical treatments for neurological and neuropsychiatric disorders.

laborious and subjective detection of DBS targets.

Building upon neurophysiological hypotheses on disease- and symptom-specific network activities, a variety of neural signal features have recently been employed to inform the DBS implantation procedure and its clinical outcome. For example, the most prominent feature in STN-DBS, the neural activity in the beta band ($\sim 13\text{--}35$ Hz) stems from strong evidence that an elevated beta power in motor regions of the cerebral cortex and basal ganglia is associated with reinforcement of the current motor state [3], a process that is pathophysiologically disturbed in the presence of rigidity and bradykinesia, two of the cardinal symptoms of PD. We, along with others, have speculated about the existence of beta-band islands, local functional neuronal organizations found in STN areas other than the dorsolateral area where one expects to find sensorimotor activity [4]. This could support the idea that spatially distributed synchronizations may be a key feature of the STN pathophysiology in PD and a possible future target for DBS. The DBS implantation procedure can also be informed by using the second major component of MERs, the intranuclear MUA. For example, we have extracted quantitative temporal trends (feature activity versus time) from MERs to generate spatial profiles (feature activity versus MER depth) of the nearby brain structures. By employing kernel depth-time interpolation (KDT) for the spatial profiles, we performed local-weighted averaging of multiple features, both spike dependent and spike independent, and integrated them into a fuzzy classifier [Figure 2(a) and (b)] [5]. The resultant distances to each cluster's centroid are visualized either offline or in an updated, pseudo-real-time approach [Figure 2(c) and (d)].

Subsequent identification of the STN via visualization of MER activity became a far easier and highly accurate task. Without stopping the procedure for careful recording and being susceptible to frequent spike overlaps among

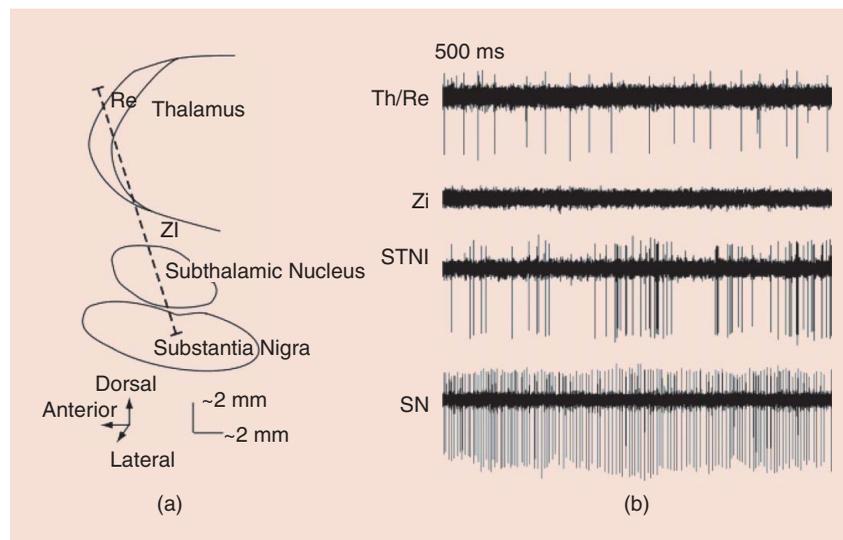


FIGURE 1. The functional targeting of the subthalamic nucleus with microelectrode recordings during DBS implantation for PD. (a) A schematic sagittal view of the typical microelectrode trajectory showing midbrain structures approximately ~ 12 mm lateral to the midline, beginning ~ 2 cm above the presumptive target. The subcortical structures along the trajectory typically include the thalamus (Th), zona incerta (ZI), subthalamic nucleus (STN), and substantia nigra (SN). (b) Distinct neurophysiological spiking and spike background patterns, corresponding to the different structures, are encountered as the electrode advances. In the example shown, the fast firing rate within SN is consistent with a typical pars reticulata (SNr) neuron.

neurons, our method paved the way for more powerful supervised learning tools and feature proliferation via, e.g., genetic methods, to further enhance the accuracy of STN detection.

Technical advances in DBS systems

There are a handful of DBS systems manufactured by Medtronic (Activa; Medtronic, Minneapolis, Minnesota), Boston Scientific (Vercise; Boston Scientific, Valencia, California), St. Jude (Infinity; St. Jude Medical, St. Paul, Minnesota), and Aleva (directStim; Aleva Neurotherapeutics, Lausanne, Switzerland). The systems comprise stimulating brain leads that target a variety of neural substrates, depending on the disorder. The implanted stimulating tip is of a quadripolar configuration with the four annular stimulating contacts clustered closely at the end. Medtronic provides the greatest detail, with each contact being 1.5 mm in height, 1.27 mm in diameter, and the spacing between contacts being either 0.5 mm or 1.5 mm, dependent upon the model. The contact materials

consist of an 80/20 platinum/iridium alloy, with the connecting wires constructed from an identical mix, coiled around a removable tungsten stylet to assist with rigidity for placement and all embedded within polyurethane for insulation, biostability, and elasticity. The insulated, nontargeted end of the stimulating lead is connected to subcutaneous extension cables running beneath the scalp and neck leading to a neurostimulator typically located subclavically on the torso. The entire system is enclosed within the

body and communicated with via radio telemetry or Bluetooth (St. Jude and Boston Scientific).

The materials and details of the DBS systems provided by Boston Scientific and Aleva Therapeutics are similar, with St. Jude being the exception utilizing stimulating contacts composed of the same platinum/iridium alloy, while connecting wires and extension contacts are composed of MP35N-LT, a nickel cobalt alloy, all embedded in ethylene tetrafluoroethylene and covered with Bionate hypo tubes (polycarbonate

We and others have speculated about the existence of beta-band islands.

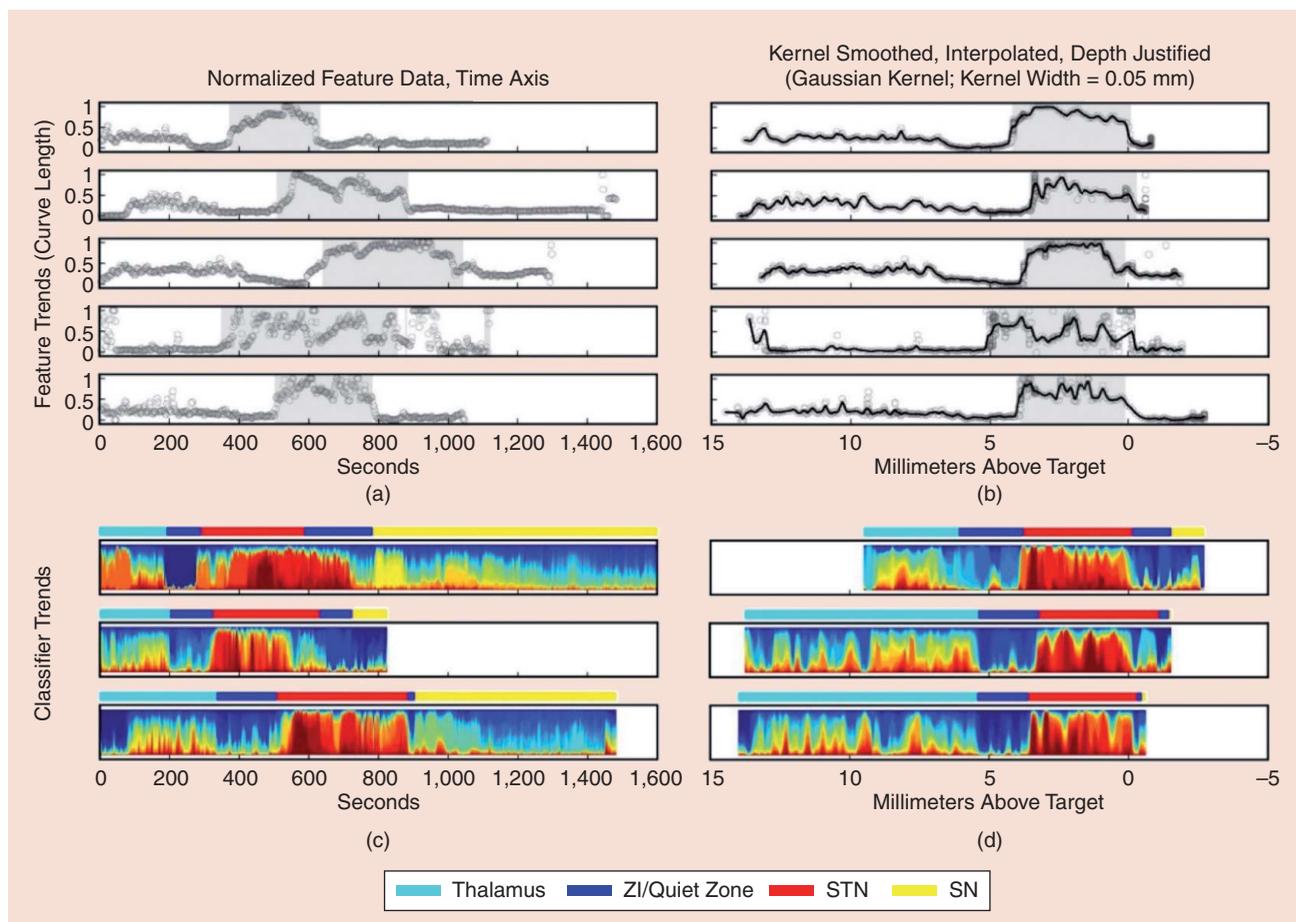


FIGURE 2. MER activity via a fuzzy clustering of multiple features for STN detection. The STN is recognizable as a red portion flanked by blue or aqua portions that represent distinct physiology obtained from white matter tracts or other neuronal structures surrounding the STN. (a) and (b) Feature trends from a single feature calculation (curve length). Open circles represent feature activity normalized to the data window length. STN boundaries are marked by gray boxes. (c) and (d) Activity maps generated via fuzzy clustering of multiple features [8]. Different subcortical structures are marked by colored bars (see legend), with the target STN in red and ventral edge located at 0 mm above target. (a) and (c) Feature trends/activity maps presented on the time-axis. (b) and (d) KDT interpolation with a Gaussian kernel (width = 0.05 mm). In (b) and (d), the thick black lines indicate the result of KDT and interpolation (with 1,000 points). For (c) and (d), the interpolation of normalized feature trends is used.

polyurethane). Medtronic systems were FDA approved in 1997 (ET) and 2002/2003 (PD), while St. Jude's Infinity system was FDA approved for both in September 2016. Boston Scientific won the approval of Conformité Européene in September 2015, and the Aleva stimulating lead system is still undergoing clinical trials.

When active, DBS systems deliver a continuous train of asymmetric biphasic square waves, either current or voltage based, whose setting of multiple contact configurations and programmable parameters of amplitude, pulse width, and frequency can be adjusted to maximize

an individual's symptom control, while minimizing adverse stimulation effects (thereby maximizing the therapeutic window).

Steering the neurostimulation

Directional current steering is offered by the Boston Scientific (Vercise) and St. Jude's (Infinity) DBS systems, with only the Infinity currently being available in the United States. Horizontal steering of the stimulation fields emitted by the two middle annular contacts of the quadripolar electrode is achieved by segmenting the annular ring into three 120° partitions that

can be individually activated or deactivated. Thus, if the stimulating electrode is placed more medially than it should, the more lateral facing segments of the split ring can be activated selectively, preventing medial spillage of the stimulation field outside the desirable target region to reduce side effects. Aleva Neurotherapeutics has also engineered a similar three-way split ring stimulating electrode, except utilizing the lower two annular contacts of the quadripolar stimulation lead. These electrodes were recently trialed intraoperatively, exhibiting greater benefit for directional over omnidirectional stimulation in 13 movement disorder DBS candidates [6].

Much of the split ring annular electrodes may salvage the benefit of DBS

Much of the split ring annular electrodes may salvage the benefit of DBS for nonoptimally placed leads.

for nonoptimally placed leads. Yet, the split ring does not provide the resolution and shaping capabilities of the 32 contact stimulating lead trialed intraoperatively by Contarino et al [7]. The 32 tessellating contact resolution allows the annular shape to slide up and down the stimulating lead in a near continuous fashion, in contrast to jumping between nonoverlapping dorsal-ventral segments. As such, the sculpting possibilities are endless, fulfilling the real potential of stimulation field shaping. Not surprisingly, the results from the intraoperative pilot indicated that directional stimulation could be increased much greater before evoking adverse events than could the annular mimicking stimulation.

Widening the neurostimulation parameter space

As attractive as physical shaping of the stimulation field may appear, real advances may alternately be available through a better understanding of programmable stimulation parameters of pulse frequency, width, and amplitude. The parameter values that are used today have been dictated by the technical limitations of the available neurostimulation devices and tuned within these limits by clinical experience. Recent studies have proposed various ways for widening the parameter space with the goal of selectively stimulating therapeutic target neurons at the lowest energy possible.

Varying temporal patterns through interleaving is unique to Medtronic's Activa family of neurostimulators and was first introduced in 2009. Here, alternating pulses are emitted from different contacts of the same stimulating lead, each with independently programmable amplitude and pulse width, but with the same interdigitated frequency. This was originally intended to allow dual regions of a target substrate along the dorsal/ventral axis of the stimulating lead to be activated, while leaving the region in between unperturbed. Thus, it was thought that multiple symptoms could be captured by multiple sites.

However, if the stimulation fields are brought in close proximity to each other, either by using adjacent contacts

or by increasing the amplitude and pulse width, then the two stimulation fields may overlap creating a region of stimulation that will receive twice the programmed frequency, in addition to nonoverlapped regions receiving the programmed frequency. As such, two-tiered frequency stimulation fields can be sculpted, allowing multiple symptoms to be captured or alternately adverse effects released by engaging multiple temporal frequencies.

While the St. Jude (Infinity) system can drive the stimulation of different leads at independent frequencies, the Boston Scientific (Vercise) DBS system is capable of programming independent frequencies on the same lead, for two active "areas," defined as any aggregate of contacts and/or contact segments. Thus, the Vercise system can create temporal "patterned" stimulation in the overlapped regions of the generated fields. The two areas would be driven at different frequencies, with the initial stagger interval between them being determined by the lagging anodal phase of the initial area pulse, ultimately resulting in doublets or triplets instead of continuous stimulation trains. Medtronic's interleaving and Boston Scientific's staggered independent frequencies are two ways to implement multiple frequency fields or patterned stimulation. However, the clinical significance of these new technical capabilities remains to be seen. If the utility of these simple temporal stimulation patterns can be clinically demonstrated then more complicated bursting capabilities could be intentionally engineered. Neurons in the brain lend themselves to bursting, why not DBS?

Future perspectives: Toward adaptive and precise neuromodulation

Many open questions on the neural underpinnings of neuromodulation are expected to be addressed by recent scien-

tific and technical achievements in DBS systems. Notwithstanding the dramatic improvement that DBS already brings to the quality of life for many patients, we are far from securing, if not defining, its maximum clinical outcome. For DBS implantation, two straightforward objectives are to provide 1) pre- and intra-operative support in localizing the DBS target area and 2) neuromarkers that depict the neurophysiological variability and, therefore, are predictive of the DBS outcome. Especially for psychiatric diseases that are typically believed to be due to brain network imbalances, DBS localization is expected to benefit from approaches that link DBS with other, noninvasive, stimulation techniques applied on the same functional networks [8]. For DBS programming, one possible objective is to step away from the stereotyped stimulation patterns that current open-loop DBS systems provide and move

toward neuromodulation that adapts at the millisecond scale, where neurons communicate.

As a surgical treatment for movement disorders, DBS has been historically delivered in an open-loop fashion where a

preprogrammed, chronic and continuous stimulation pattern could not avert suboptimal clinical outcomes. Leveraging the technical advances in new DBS devices, clinical studies show that a closed-loop DBS (CL-DBS) system is realizable. What still seems elusive is the driving signals for such systems, i.e., the neural signals and their features, that are informative enough to control the online real-time adaptation of the neuromodulator. One might argue that, for the current technology, the best control signal is the LFP, or some component of it. The reason is that LFP represents the neural information integrated over a larger area compared to the multiunit activity and, therefore, presumably carries more information about the cardinal symptoms of the disease and can account for intersubject

Many open questions on the neural underpinnings of neuromodulation are expected to be addressed by recent scientific and technical achievements in DBS systems.

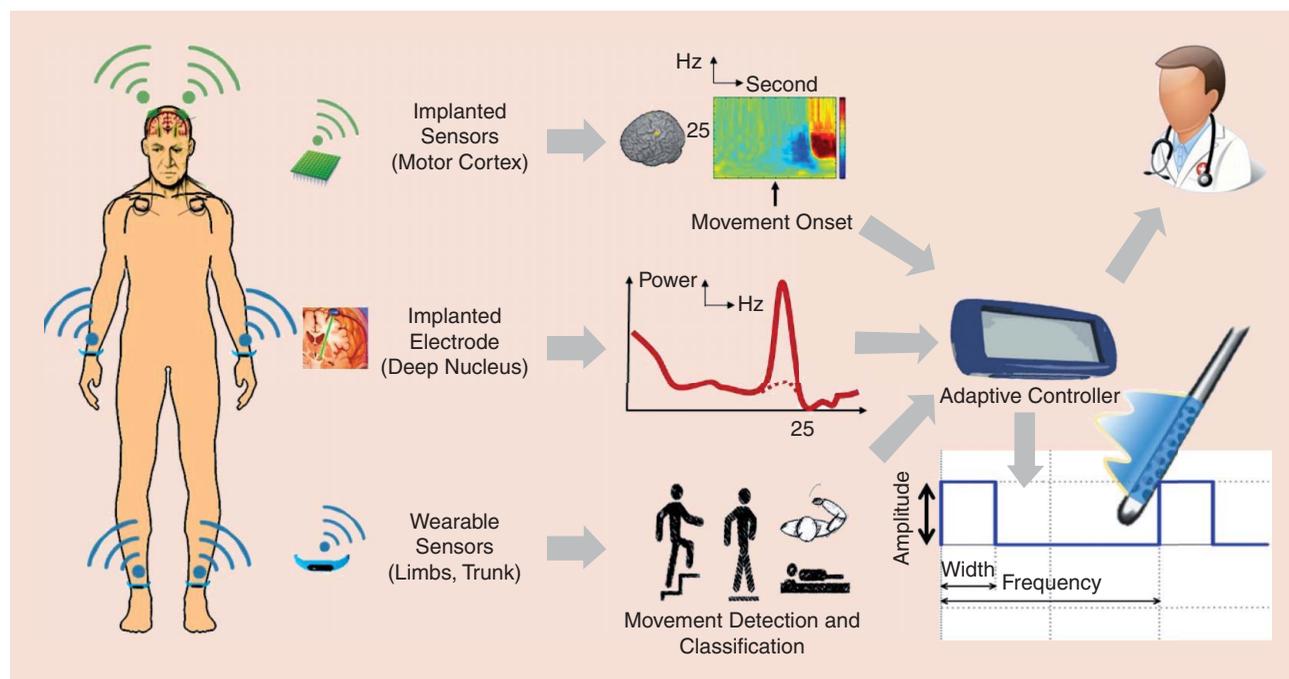


FIGURE 3. Closed-loop DBS controlled by neural signals acquired from multiple modalities across the spatial and temporal scales of brain function. A DBS controller adapts in real-time the therapeutic parameters (current distribution, frequency, amplitude, width) based on an algorithm that is driven by combining features extracted from 1) wireless, batteryless sensors implanted on cortical areas (e.g., motor cortex), 2) single and multineuron recordings (through the DBS implanted electrodes), and 3) wearable sensors identifying patterns of movement.

variability, thus enabling sufficient personalization. It is not a coincidence that a specific feature of the LFP, the beta-band activity of STN neurons, is currently under intense testing on whether it can become an effective programming biomarker for CL-DBS for PD [4], [9]. The increased beta band within the STN may also represent a greater coherence and phase locking of this oscillation across multiple basal ganglia structures as part of the underlying pathophysiology of the system [10]. Incorporating other LFP oscillations such as theta frequency afferent from the medial prefrontal cortex to the STN [11] or gamma frequencies, found to correlate with PD symptoms, seem to be a natural next step. Moving beyond the MERs, one should consider the use of other neural as well as behavioral signals transmitted wirelessly to a wearable information processing platform. A potential helpful expansion of a DBS system could employ the fusion of neural and other

A potential helpful expansion of a DBS system could employ the fusion of neural and other signals acquired via multiple modalities.

signals acquired via multiple modalities, including wearable and implantable sensors (Figure 3). Implanting wireless sensors on the motor cortex and coupling the neural information with behavioral signals acquired through wearable devices that classify movement patterns could provide new information pathways (e.g., movement-triggered cortical oscillations such as beta-band rebound or mu-alpha suppression) toward controlling a CL-DBS system by integrating features from multiple modalities. We have shown earlier that employing a small number of neurophysiologically interpretable features inside the STN can predict, separately for each patient, the behavioral outcome of STN-DBS. The neurophysiological basis of using implanted wireless sensors of brain activity in the motor cortex stems from the fact that stimulating STN neurons can cause antidromic activation of the hyperdirect pathway, which consists of axon collaterals of pyramidal neurons

in the motor cortex. Stimulating these axons within the STN is associated with changes in motor cortical activity, possibly masking or desynchronizing pathologically enhanced beta-band oscillations within the basal ganglia-thalamocortical network. Therefore, the goal for a CL-DBS system should be to maximize the stimulation of these target neurons while minimizing unintended activation of nontarget neurons such as corticospinal or corticobulbar fibers within the internal capsule, which may cause speech, walking, or fine motor skill impairments. Overall, a CL-DBS system will not only secure the clinical effectiveness but also minimize the potential for serious complications and side effects.

Nevertheless, even if these objectives are met, the currently available DBS systems are presumed to modulate more cells than those affected by the disease, which could sometimes lead to side effects. Thanks to recent advances in neurosciences and signal processing, we are getting close to the development of electroceuticals, systems aiming to modulate the spike activity of individual and functional

groups of neurons in adaptive ways that are fully compatible with the biological function [12]. To this end, we will need a better mapping of the neural circuits associated with the treated pathophysiology; at the signal level, we will need better decoders of the neural language associated with the pathophysiological states and more precise therapeutic patterns of electrical impulses targeting the rate, even the timing of spikes. Generating such adaptive and precise neuromodulators will require a multidisciplinary effort: the development of neuromorphic circuits for real-time spike processing will translate the biological understanding of what is happening at the neural level in health and disease. That said, we shouldn't underestimate the complexity of such an endeavor that could result in another big data mining problem, this time at the neural level. Such problems can only be approached synergistically; to achieve this, we need initiatives that bring together scientists and engineers, the most prominent of which is the yearly workshop on neuro-modulation organized by the Institute of Engineering in Medicine at the University of Minnesota.

Conclusions

Recent advances in basic and clinical neuroscience have helped us understand which should be the target neurons for a particular DBS indication and which neural elements within the stimulation volume rather contribute to adverse effects of stimulation. Progress in medical technology has allowed the development of new DBS devices with unprecedented technical abilities that now offer a more refined, in time and space, neuro-modulation. Ongoing computational analyses are proposing neurophysiologically optimized solutions for DBS while removing a significant burden for advanced clinical experience and repeated intra- and postoperative test-

ing of the patient response. A tight interweaving of the multidisciplinary advances will allow the validation of neurophysiological concepts of neurostimulation in clinical practice and translate DBS, from a complex and poorly standardized therapy where treatment failures are not uncommon, to a flexible intervention, tailored to each patient's symptoms and neuropathophysiology. This direction merits further research.

Authors

Konstantinos P. Michmizos (konstantinos.michmizos@cs.rutgers.edu) is an assistant professor of computer science and the director of the Computational Brain Lab at Rutgers University, Piscataway, New Jersey.

Blerita Lindqvist (b.lindqvist@rutgers.edu) is a Ph.D. student in the Computational Brain Lab, Department of Computer Science, Rutgers University, Piscataway, New Jersey.

Stephen Wong (wongst@rwjms.rutgers.edu) is an associate professor of neurology at Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey.

Eric L. Hargreaves (hargreer@rwjms.rutgers.edu) is an instructor of neurosurgery and the director of clinical deep brain stimulation research at Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey.

Konstantinos Psychas (kpsychas@biosim.ntua.gr) is a graduate of the National Technical University of Athens and a Ph.D. student at Columbia University, New York.

Georgios D. Mitsis (georgios.mitsis@mcgill.ca) is an assistant professor of bioengineering at McGill University, Montreal, Canada.

Shabbar F. Danish (shabbar.danish@rutgers.edu) is the chief of neurosurgery at Rutgers Cancer Institute and the director of Stereotactic and Functional

Neurosurgery and Gamma Knife Center at Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey.

Konstantina S. Nikita (knikita@ece.ntua.gr) is a professor of electrical and computer engineering at the National Technical University of Athens, Greece.

References

- [1] A. L. Benabid, "Deep brain stimulation for Parkinson's disease," *Curr. Opin. Neurobiol.*, vol. 13, no. 6, pp. 696–706, 2003.
- [2] K. P. Michmizos, D. Sakas, and K. S. Nikita, "Prediction of the timing and the rhythm of the parkinsonian subthalamic nucleus neural spikes using the local field potentials," *IEEE Trans. Inform. Technol. Biomed.*, vol. 16, no. 2, pp. 190–197, Mar. 2012.
- [3] A. K. Engel and P. Fries, "Beta-band oscillations—signalling the status quo?" *Curr. Opin. Neurobiol.*, vol. 20, no. 2, pp. 156–165, Apr. 2010.
- [4] K. P. Michmizos, P. Frangou, P. Stathis, D. Sakas, and K. S. Nikita, "Beta-band frequency peaks inside the subthalamic nucleus as a biomarker for motor improvement after deep brain stimulation in Parkinson's disease," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 174–180, Jan. 2015.
- [5] S. Wong, G. Baltuch, J. Jaggi, and S. Danish, "Functional localization and visualization of the subthalamic nucleus from microelectrode recordings acquired during DBS surgery with unsupervised machine learning," *J. Neural Eng.*, vol. 6, Mar. 2009.
- [6] A. L. Green and T. Z. Aziz, "Steering technology for deep brain stimulation," *Brain*, vol. 137, no. 7, pp. 1854–1856, May 2014.
- [7] M. F. Contarino, L. J. Bour, R. Verhagen, M. A. Lourens, R. M. de Bie, P. van den Munckhof, and P. R. Schuurman, "Directional steering: A novel approach to deep brain stimulation," *Neurology*, vol. 83, pp. 1163–1169, Sept. 2014.
- [8] M. D. Fox, R. L. Buckner, H. Liu, M. M. Chakravarty, A. M. Lozano, and A. Pascual-Leone, "Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 111, pp. E4367–E4375, Sept. 2014.
- [9] L. A. Johnson, S. D. Nebeck, A. Muralidharan, M. D. Johnson, K. B. Baker, and J. L. Vitek, "Closed-loop deep brain stimulation effects on Parkinsonian motor symptoms in a non-human primate—Is beta enough?" *Brain Stimul.*, vol. 9, no. 6, pp. 892–896, 2016.
- [10] H. Cagnan, E. P. Duff, and P. Brown, "The relative phases of basal ganglia activities dynamically shape effective connectivity in Parkinson's disease," *Brain*, vol. 138, no. 6, pp. 1667–1678, 2015.
- [11] B. Zavala, H. Tan, K. Ashkan, T. Foltyniec, P. Limousin, L. Zrinzo, K. Zaghoul, and P. Brown, "Human subthalamic nucleus-medial frontal cortex theta phase coherence is involved in conflict and error related cortical monitoring," *NeuroImage*, vol. 137, pp. 178–187, May 2016.
- [12] K. Famm, B. Litt, K. J. Tracey, E. S. Boyden, and M. Slaoui, "Drug discovery: A jump-start for electroceuticals," *Nature*, vol. 496, no. 7444, pp. 159–161, 2013.

ADVERTISING & SALES



© GRAPHIC STOCK

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*.

COMPANY	PAGE NUMBER	WEBSITE	PHONE
Mathworks	CVR 4	www.mathworks.com/wireless	+1 508 647 7040

445 Hoes Lane, Piscataway, NJ 08854

IEEE SIGNAL PROCESSING MAGAZINE REPRESENTATIVE

Mark David
 Director, Business Development — Media & Advertising
 Phone: +1 732 465 6473
 Fax: +1 732 981 1855
m.david@ieee.org

Digital Object Identifier 10.1109/MSP.2017.2660358

IEEE Journal of Selected Topics in Signal Processing (JSTSP)

Recent Special Issues

- ❖ March 2017 – Time/Frequency Modulated Array Signal Processing
- ❖ February 2017 – Measuring Quality of Experience for Media Technologies and Services
- ❖ November 2016 – Exploiting Interference for Energy-Efficient & Secure Communications
- ❖ October 2016 – Advanced Signal Processing for Brain Networks
- ❖ September 2016 – Financial Signal Processing and Learning for Electronic Trading
- ❖ August 2016 – Person-Centered Signal Processing for Health Technologies
- ❖ June 2016 – Structured Matrices in Signal and Data Processing
- ❖ April 2016 – Signal Processing for Millimeter Wave Communications

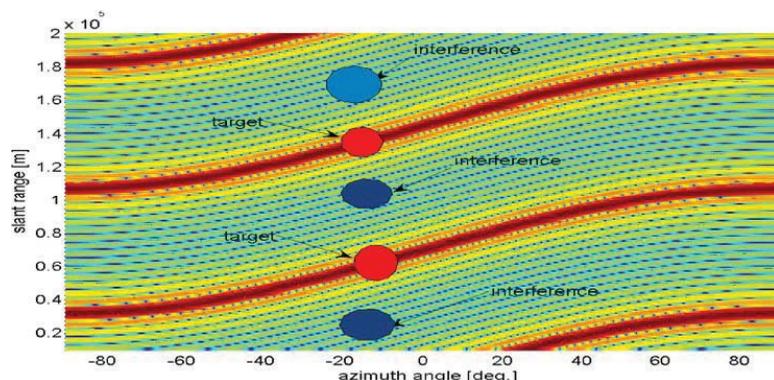


Image credit: <https://doi.org/10.1109/JSTSP.2016.2627182>

Digital Object Identifier 10.1109/MSP.2017.2673079

IN THE SPOTLIGHT *(continued from page 124)*

can have an impact. Our algorithms and our technical areas are precisely the tools that are needed to advance vehicular systems. We have been conducting research in a broad range of fields, such as radar, computer vision, and statistical signal processing that can allow us to revolutionize the field, to develop the necessary innovations that will make automotive systems reliable to the tenth decimal place. Here we take a look at this problem and present four promising ways signal processing and data analytics can have an impact on the challenges surrounding automated vehicles.

Properly assess and utilize data for safety decisions

Cars currently contain roughly 100 sensors, and future automobiles will likely be deployed with significantly more, including accelerometers (for impact detection and motion measurements), pressure sensors (for air intake control, monitoring fuel consumption, tire conditions), temperature sensors (to monitor and control engine conditions, fuel temperature, passenger compartment temperature), and phase sensors (camshaft/crankshaft phase sensors for motor control, and gear shaft speed for transmission control). Angular rate sensors monitor the roll, pitch, and yaw of a vehicle, which informs dynamic control systems, automatic distance control, and navigation systems. Angular and position sensors monitor the position of gear levers, steering wheel angle, and mirror positioning. Radar, lidar, and camera sensors are used to facilitate new applications, such as blind spot monitoring, lane-departure warning, and automated driving.

These sensors provide the abundance of data that can serve as corroborating evidence to fix malfunctions, back-solve and determine a mistake is about to be made from using a single sensor type alone, and correct false data injected by those trying to hack our vehicles. When properly utilized, this

wealth of data is the avenue to safety and robustness.

Merge multiple types of imaging sensors for fast object recognition

It's evident that the Tesla crash videos recorded by Autopilot weren't under ideal lighting conditions. Background objects blended into vehicles that needed to be recognized, making it difficult for any computer to process correctly. This was amplified by the short time allowed to "lock on" given the speed of the vehicle and the imminent crash. Multiple sensor types used in conjunction could have helped. Radar or lidar would not have been susceptible to the same difficulties the camera-based system likely encountered.

Tesla has since reevaluated its strategy for Autopilot, including the possibility of using radar in place of the camera, and two things are clear: the choice of a radar system is meant to avoid the environmental hurdles that arise with visual-based systems, and Tesla has collected a large amount of radar data that serves as the basis of its new Autopilot system.

Though likely an improvement, switching to a single type of sensor isn't likely to solve all the problems that will arise in automating vehicles. In fact, while radar can cope with lighting-based challenges, numerous studies suggest lidar systems are superior in terms of tracking accuracy. While lidar systems suffer degradation in conditions with fog, cameras offer the ability to recognize finer details associated with objects (such as license plate information). In fact, cameras support the accurate assessment of the visibility distance (notably fog), which could be used to inform the driver that vehicular assistance services aren't available or are experiencing degraded quality of service because fog is affecting the visibility of road lanes and other vehicles. Data fusion and extracting hidden correlation between sensor types is at the heart of modern signal processing. Merging radar,

lidar, and visual systems into fast and robust object recognition and tracking algorithms is an exciting opportunity where signal processors can contribute.

Share data between vehicles to correct miscalculations or other errors

When considering future vehicular applications, we should recognize other sensor types will be available and can provide valuable knowledge, like weather conditions, road friction coefficients, or road slopes. Road slope information is useful for coordinating braking among several vehicles since slope is related to the potential for a vehicle to accelerate or decelerate. Data sharing between vehicles and cloud-based computing services opens up many other possibilities to improve vehicle safety. Data shared between vehicles will allow signal processing algorithms running on each vehicle to gather the conditions that may be experienced by other nearby vehicles.

Furthermore, data measured by the multitude of sensors, whether from within a single vehicle or across several, can be used to correct malfunctioning or poorly calibrated sensors. Currently, vehicular sensors are recalibrated by bringing a vehicle to a certified garage to update or replace the sensor. By using the distributed nature of the vehicular setting—in which there are numerous vehicles frequently making data measurements correlated across many dimensions—it becomes possible to report this data to cloud servers that would perform large-scale data analytics to accurately identify the corrections needed.

Understand human driving behavior through signal processing advancements

Going beyond the technical aspects, what is forgotten is that transportation also serves as a complex social fabric by which we interact with each other. This merging of "cyberphysical" with

social makes automating vehicles an ultrahard problem to tackle. We observe cues from vehicles while we drive that suggest we should drive more cautiously (e.g., a pedestrian looking at his smartphone while walking toward an intersection) or even avoid certain driving scenarios (e.g., an overzealous driver swerving through lanes). While it's unreasonable to expect that self-driving cars could make the same social observations that we make as humans, what we can expect is that technology will assist us in being as aware and informed as possible. Already there have been advancements made by the signal processing community to estimate driver distraction using in-vehicle sensors and cue the driver to focus on the road. However, there are many other opportunities for signal processing engineers to analyze human behavior data associated with driving,

which will be essential for improving driver and pedestrian safety.

The future of vehicular systems is data and sensor driven. Vehicles will become increasingly networked and outfitted with sensors and share their data with a variety of in-vehicle and cloud-based computing services. The societal benefits associated with improved vehicular systems range from energy efficiency resulting from swarm driving to the potential for saving many lives should the technology mature. While this future is exciting, engineers, researchers, and technologists must quickly act to develop the new signal and information processing innovations required to make future vehicular systems safe.

Author note

Some parts of this article originally appeared on Robotic Tips website; <http://www.robotictips.com>.

Author

Wade Trappe (trappe@winlab.rutgers.edu) received his B.A. degree in mathematics from the University of Texas at Austin and his Ph.D. degree in applied mathematics and scientific computing from the University of Maryland. He is an IEEE Fellow, and a professor of electrical and computer engineering at Rutgers University, New Brunswick, New Jersey.

References

- [1] D. Ngo. (14, Sept. 2016). Tragic Tesla crashes in China. [Online]. Available: www.cnet.com/news/dash-cam-showed-fatal-tesla-crash-in-china
- [2] D Shepardson. (26 July 2016). Tragic Tesla crashes in the U.S. [Online]. Available: www.reuters.com/article/us-tesla-autopilot-idUSKCN1062CT
- [3] A Webb. (19 July 2016). Cybersecurity is biggest risk of autonomous cars: Survey finds. [Online]. Available: www.bloomberg.com/news/articles/2016-07-19/cybersecurity-is-biggest-risk-of-autonomous-cars-survey-finds

SP

READER'S CHOICE (continued from page 13)

Tensor Decompositions for Signal Processing Applications: From Two-Way to Multiway Component Analysis

Cichocki, A.; Mandic, D.; Lathauwer, L. De; Zhou, G.; Zhao, Q.; Caiafa, C.; Phan, H.A.

The authors have shown that tensor decompositions are a good match for exploratory analysis of multifaceted data sets and have illustrated their applications in multisensor and multimodal signal processing. Their emphasis has shown that tensor decompositions and multilinear algebra open up completely new possibilities for component analysis, as compared with the flat view of standard two-way methods.

March 2015

Euclidean Distance Matrices: Essential Theory, Algorithms, and Applications

Dokmanic, I.; Parhizkar, R.; Ranieri, J.; Vetterli, M.

This article reviews the fundamental properties of Euclidean distance matrices (EDMs) and shows how the various EDM properties can be used to design algorithms for completing and denoising distance data. Some directions are given for further research.

November 2015

Bayesian Machine Learning: EEG/MEG Signal Processing Measurements

Wu, W.; Nagarajan, S.; Chen, Z.

To review recent advances and to foster new research ideas, this article provides a tutorial on several important emerging

Bayesian machine learning research topics in electroencephalography (EEG)/magnetoencephalography (MEG) signal processing and presents representative examples in EEG/MEG applications.

January 2016

Compressive Covariance Sensing: Structure-Based Compressive Sensing Beyond Sparsity

Romero, D.; Ariananda, D.D.; Tian, Z.; Leus, G.

This article is concerned with the reconstruction of second-order statistics, such as covariance and power spectrum, instead of the reconstruction of signals in compressed sensing, even in the absence of sparsity priors.

January 2016

SP

DATES AHEAD

Please send calendar submissions to:
Dates Ahead, Att: Jessica Barragué, E-mail: j.barrague@ieee.org

2017

APRIL

IEEE International Symposium on Biomedical Imaging (ISBI)

18–21 April, Melbourne, Australia.

General Chairs: Olivier Salvado and Gary Egan

URL: <http://biomedicalimaging.org/2017/>

16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)

18–21 April 2017, Pittsburgh, Pennsylvania, USA.

General Chair: Pei Zhang

URL: <http://ipsn.acm.org/2017/>

MAY

IEEE Radar Conference (RADARCONF)

8–12 May, Seattle, Washington, USA.

General Chair: Daniel J. Segó

URL: <http://www.radarconf17.org/>

JULY

18th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)

3–6 July, Hokkaido, Japan.

General Chairs: Yasutaka Ogawa, Wei Yu, and Fumiyuki Adachi

URL: <http://www.spawc2017.org/>

IEEE International Conference on Multimedia and Expo (ICME)

10–14 July, Hong Kong, China.

General Chairs: Jörn Ostermann and Kenneth K.M. Lam

URL: <http://www.icme2017.org/>



The IEEE International Symposium on Biomedical Imaging will be held 18–21 April in Melbourne, Australia.

AUGUST

25th European Signal Processing Conference (EUSIPCO)

28 August–2 September, Kos Island, Greece.

General Chairs: Petros Maragos and Sergios Theodoridis

URL: www.eusipco2017.org

14th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)

29 August–1 September, Lecce, Italy.

General Chairs: Cosimo Distanto and Larry S. Davis

URL: www.avss2017.org

SEPTEMBER

IEEE International Conference on Image Processing (ICIP)

17–20 September, Beijing, China.

General Chairs: Xingang Lin, Anthony Vetro, and Min Wu

URL: <http://2017.ieeeicip.org/>

OCTOBER

IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)

15–18 October, New Paltz, New York, USA.

General Chairs: Patrick A. Naylor and Meinard Müller

URL: <http://www.waspaa.com/>

19th IEEE International Workshop on Multimedia Signal Processing (MMSP)

16–18 October, London-Luton, United Kingdom.

General Chairs: Vladan Velisavljevic, Vladimir Stankovic, and Zixiang Xiong

URL: <http://mmsp2017.eec.strath.ac.uk/>

NOVEMBER

Fifth IEEE Global Conference on Signal and Information Processing (GlobalSIP)

14–16 November 2017, Montreal, Canada.

General Cochairs: Warren Gross and Kostas Plataniotis

URL: <http://2017.ieeeglobalsip.org>

DECEMBER

Seventh IEEE Conference of the Sensor Signal Processing for Defence (SSPD)

6–7 December, Edinburgh, Great Britain.

General Chairs: Mike Davies, Jonathon Chambers, and Paul Thomas

URL: www.sspd.eng.ed.ac.uk/

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

No Need for Speed: More Signal Processing Innovation Is Required Before Adopting Automated Vehicles

We spend a considerable amount of time driving—to work, to home, for recreation, for travel. This love for the automobile is on the precipice of becoming a worldwide phenomenon. While new markets, like China, have opened up in the past decade, the need for the automotive industry to find new customers, combined with an array of emerging technologies that will make driving easier, will allow cars to reach markets that never before had access to driving. Likewise, the promise of new capabilities that automate and enhance the safety of the driving experience will guarantee that existing drivers return to purchase the latest model.

People will be able to go to a terminal and rent a vehicle preprogrammed to take them to a specified destination. A driver will be able to disengage from actual driving to read the newspaper while the car carries out his or her daily commute to and from work—all the while, cars will seamlessly coordinate to ensure the safety of their passengers.

And why will this happen? Because people like the freedom and excitement that comes with “getting in the car and going places!”

This promise is very enticing, but society shouldn't rush headlong into adopting automated vehicular systems. Transportation systems are very complex systems with many interacting



IMAGE LICENSED BY GRAPHIC STOCK

pieces—anyone who has driven through rush hour in any of the world's urban centers will attest to this!

Recently, there have been several stories in the news that have emphasized the challenges facing the automation of vehicular systems. Two tragic crashes that occurred in Tesla vehicles (one in China [1] and one in the United States [2]) highlight the serious life-and-death consequences associated with malfunctions or miscalculations that can occur with vehicular systems. Meanwhile, other news stories [3] have given us insight into what can go wrong should these systems and their data come in the cross-hairs of cybercriminals.

What can be done about this? Perhaps first and foremost is to slow our rush to “remove the human” from the equation. Officially, Tesla's Autopilot was meant to assist the driver, not replace the driver. Humans are still essential to driving, and, quite frankly, it should be that way for some time until the technology matures. Let's be clear, though: this does not mean that we should slow down innovation but rather that we should work harder to provide even more innovation that can make its way into vehicles!

It is toward this vision where we, as signal processors and data analysts,

(continued on page 121)



IEEE Signal Processing Society Video and Image Processing (VIP) Cup: *Traffic Sign Detection under Challenging Conditions*

The IEEE Signal Processing Society announces the first edition of the Signal Processing Society Video and Image Processing (VIP) Cup: *traffic sign detection under challenging conditions*.



Robust and reliable traffic sign detection is necessary to bring autonomous vehicles onto our roads. State of the art traffic sign detection algorithms in the literature successfully perform the task over existing databases that mostly lack realistic road conditions. This competition focuses on detecting such traffic signs under challenging conditions.

To facilitate such task and competition, we introduce a novel video dataset that contains a variety of road conditions. In such video sequences, we vary the type and the level of the challenging conditions including a range of lighting conditions, blur, haze, rain and snow levels. The goal of this challenge is to implement traffic sign detection algorithms that can robustly perform under such challenging environmental conditions.

Any eligible team can participate in the competition, whose detailed guidelines and dataset are planned to be released on **March 15, 2017** and participating teams should complete their submission by **July 1, 2017**. The three best teams are selected and announced by **August 1, 2017**. Three finalist teams will be judged at ICIP 2017, which will be held **September 17-20, 2017**. In addition to algorithmic performances, demonstration and presentation performances will also affect the final ranking.

The champion team will receive a **grand prize** of \$5,000. The first and the second runner-up will receive a prize of \$2,500 and \$1,500, respectively, in addition to travel grants and complimentary conference registrations. Each finalist team invited to ICIP 2017 will receive travel grant supported by the SPS on a reimbursement basis. A team member is offered up to \$1,200 for continental travel, or \$1,700 for intercontinental travel. A maximum of three members per team will be eligible for travel support.

For more details, please refer to the main web pages of the VIP Cup:

<http://signalprocessingsociety.org/get-involved/video-image-processing-cup>

<https://ghassanalregib.com/vip-cup/>

Digital Object Identifier 10.1109/MSP.2017.2663738

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

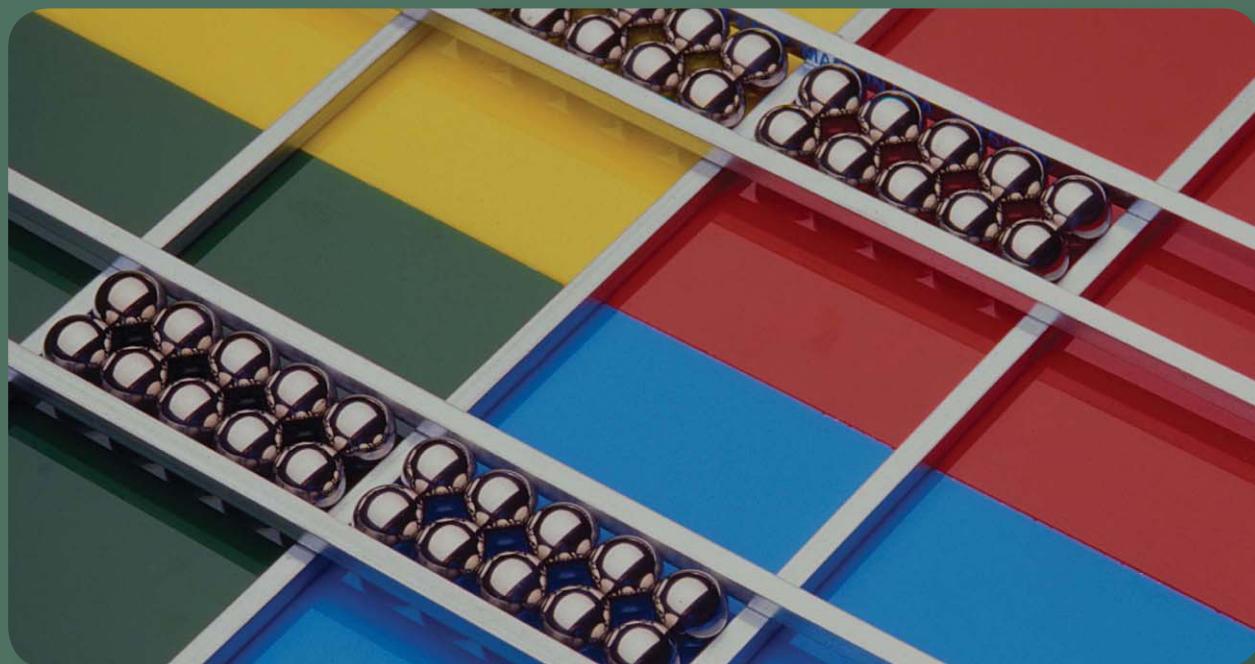


IEEE SIGNAL PROCESSING SOCIETY

Content Gazette

MARCH 2017

ISSN 2167-5023



T-SP March 1 Vol. 65 #5

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7820229>

T-SP March 15 Vol. 65 #6

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7840085>

T-ASLP February Vol. 25 #2

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7827191>

T-IP February Vol. 26 #1

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7815513>

T-IFS February Vol. 12 #2

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7827177>

T-IFS March Vol. 12 #1

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7833219>

T-MM February Vol. 19 #2

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7820238>

T-SPL February Vol. 24 #2

<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7842753>

http://www.ieee.org/publications_standards/publications/authors/author_ethics.html



IEEE ASRU 2017

Okinawa, Japan,
December 16-20, 2017



General Chairs:

John R. Hershey, MERL
Tomohiro Nakatani, NTT

Important Dates:

Paper Submission:

June 29, 2017

Paper Notification:

August 31, 2017

Early Registration Period:

August 31 - Oct 5, 2017

Camera Ready Deadline:

Sept 21, 2017

More Information:

<http://asru2017.org>

info@asru2017.org

IEEE Automatic Speech Recognition and Understanding Workshop

The biennial IEEE ASRU workshop has a tradition of bringing together researchers from academia and industry in an intimate and collegial setting to discuss problems of common interest in automatic speech recognition, understanding, and related fields of research. The workshop includes keynotes, invited talks, poster sessions and will also feature challenge tasks, panel discussions, and demo sessions.

We invite papers in all areas of spoken language processing, with emphasis placed on the following topics:

- Automatic speech recognition
- ASR in adverse environments
- New applications of ASR
- Speech-to-speech translation
- Spoken document retrieval
- Multilingual language processing
- Spoken language understanding
- Spoken dialog systems
- Text-to-speech systems



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



MARCH 1, 2017

VOLUME 65

NUMBER 5

ITPREP

(ISSN 1053-587X)

REGULAR PAPERS

Local Canonical Correlation Analysis for Nonlinear Common Variables Discovery http://dx.doi.org/10.1109/TSP.2016.2628348	1101
..... <i>O. Yair and R. Talmon</i>	
Space-Time Transmit Code and Receive Filter Design for Colocated MIMO Radar http://dx.doi.org/10.1109/TSP.2016.2633242	1116
..... <i>G. Cui, X. Yu, V. Carotenuto, and L. Kong</i>	
Subcarrier-Pairing-Based Resource Optimization for OFDM Wireless Powered Relay Transmissions With Time Switching Scheme http://dx.doi.org/10.1109/TSP.2016.2628351	1130
..... <i>Y. Shen, X. Huang, K. S. Kwak, B. Yang, and S. Wang</i>	
Blind Identification of Graph Filters http://dx.doi.org/10.1109/TSP.2016.2628343	1146
..... <i>S. Segarra, G. Mateos, A. G. Marques, and A. Ribeiro</i>	
Phase Desynchronization: A New Approach and Theory Using Pulse-Based Interaction http://dx.doi.org/10.1109/TSP.2016.2633246	1160
..... <i>T. Anglea and Y. Wang</i>	



Decentralized Estimation With Dependent Gaussian Observations http://dx.doi.org/10.1109/TSP.2016.2631463	<i>F. Peng and B. Chen</i>	1172
Hypothesis Testing With Dependent Observations http://dx.doi.org/10.1109/TSP.2016.2633243	<i>S. Sobhiyeh and M. Naraghi-Pour</i>	1183
Robust Multilinear Tensor Rank Estimation Using Higher Order Singular Value Decomposition and Information Criteria http://dx.doi.org/10.1109/TSP.2016.2620965	<i>T. Yokota, N. Lee, and A. Cichocki</i>	1196
Closed-Form and Near Closed-Form Solutions for TDOA-Based Joint Source and Sensor Localization http://dx.doi.org/10.1109/TSP.2016.2633784	<i>T.-K. Le and N. Ono</i>	1207
On the Statistical Invariance for Adaptive Radar Detection in Partially Homogeneous Disturbance Plus Structured Interference http://dx.doi.org/10.1109/TSP.2016.2620115	<i>D. Ciuonzo, A. De Maio, and D. Orlando</i>	1222
Rapid Exact Signal Scanning With Deep Convolutional Neural Networks http://dx.doi.org/10.1109/TSP.2016.2631454	<i>M. Thom and F. Gritschneider</i>	1235
Multidimensional Nonseparable Oversampled Lapped Transforms: Theory and Design http://dx.doi.org/10.1109/TSP.2016.2633240	<i>S. Muramatsu, K. Furuya, and N. Yuki</i>	1251
Group-Blind Detection for Uplink of Massive MIMO Systems http://dx.doi.org/10.1109/TSP.2016.2620958	<i>G. C. Ferrante, G. Geraci, T. Q. S. Quek, and M. Z. Win</i>	1272
Adaptive Statistical Bayesian MMSE Channel Estimation for Visible Light Communication http://dx.doi.org/10.1109/TSP.2016.2630036 ..	<i>X. Chen and M. Jiang</i>	1287
A Resilient Approach to Distributed Filter Design for Time-Varying Systems Under Stochastic Nonlinearities and Sensor Degradation http://dx.doi.org/10.1109/TSP.2016.2634541	<i>Q. Liu, Z. Wang, X. He, G. Ghinea, and F. E. Alsaadi</i>	1300
Sparse Approximation of a Kernel Mean http://dx.doi.org/10.1109/TSP.2016.2628353	<i>E. C. Cortés and C. Scott</i>	1310
The CPHD Filter With Target Spawning http://dx.doi.org/10.1109/TSP.2016.2597126	<i>D. S. Bryant, E. D. Delande, S. Gehly, J. Houssineau, D. E. Clark, and B. A. Jones</i>	1324
Subspace Selection for Projection Maximization With Matroid Constraints http://dx.doi.org/10.1109/TSP.2016.2634544	<i>Z. Zhang, Y. Wang, E. K. P. Chong, and A. Pezeshki</i>	1339
Multitask Additive Models With Shared Transfer Functions Based on Dictionary Learning http://dx.doi.org/10.1109/TSP.2016.2634546	<i>A. Fawzi, M. Sinn, and P. Frossard</i>	1352

COMMENTS AND REPLIES

Comments on “Distributed Identification of the Most Critical Node for Average Consensus” http://dx.doi.org/10.1109/TSP.2016.2631442 ..	<i>A. Bertrand</i>	1265
Reply to “Comments on ‘Distributed Identification of the Most Critical Node for Average Consensus’” http://dx.doi.org/10.1109/TSP.2016.2631460	<i>H. Liu, X. Cao, J. He, P. Cheng, C. Li, J. Chen, and Y. Sun</i>	1268

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



MARCH 15, 2017

VOLUME 65

NUMBER 6

ITPREP

(ISSN 1053-587X)

REGULAR PAPERS

A Sharp Condition for Exact Support Recovery With Orthogonal Matching Pursuit http://dx.doi.org/10.1109/TSP.2016.2634550	<i>J. Wen, Z. Zhou, J. Wang, X. Tang, and Q. Mo</i>	1370
Optimal Power Allocation for Average Detection Probability Criterion Over Flat Fading Channels http://dx.doi.org/10.1109/TSP.2016.2634552	<i>S. Sarıtaş, B. Dulek, A. D. Sezer, S. Gezici, and S. Yüksel</i>	1383
Robust Tomlinson–Harashima Precoding With Gaussian Uncertainties for SWIPT in MIMO Broadcast Channels http://dx.doi.org/10.1109/TSP.2016.2639461	<i>Q. Li, Q. Zhang, and J. Qin</i>	1399



Compressive Sensing (CS) Assisted Low-Complexity BeamSpace Hybrid Precoding for Millimeter-Wave MIMO Systems http://dx.doi.org/10.1109/TSP.2016.2641379	<i>C.-H. Chen, C.-R. Tsai, Y.-H. Liu, W.-L. Hung, and A.-Y. Wu</i>	1412
A Frequency Domain Test for Propriety of Complex-Valued Vector Time Series http://dx.doi.org/10.1109/TSP.2016.2639459	<i>S. Chandna and A. T. Walden</i>	1425
Compressive Interference Mitigation and Data Recovery in Cloud Radio Access Networks With Limited Fronthaul http://dx.doi.org/10.1109/TSP.2016.2639472	<i>J. Liu, A. Liu, and V. K. N. Lau</i>	1437
TDOA Positioning Irrespective of Source Range http://dx.doi.org/10.1109/TSP.2016.2630030	<i>Y. Wang and K. C. Ho</i>	1447
Received Power-Based Channel Estimation for Energy Beamforming in Multiple-Antenna RF Energy Transfer System http://dx.doi.org/10.1109/TSP.2016.2637320	<i>K. W. Choi, D. I. Kim, and M. Y. Chung</i>	1461
Stability Analysis of Closed-Loop Algorithms for Cancelling Nonlinear Memory Effects http://dx.doi.org/10.1109/TSP.2016.2631462	<i>Y.-Y. Xiao</i>	1477
DOA Estimation With Enhanced DOFs by Exploiting Cyclostationarity http://dx.doi.org/10.1109/TSP.2016.2645542	<i>J. Liu, Y. Lu, Y. Zhang, and W. Wang</i>	1486
Towards Robustness in Residue Number Systems http://dx.doi.org/10.1109/TSP.2016.2641398	<i>L. Xiao, X.-G. Xia, and H. Huo</i>	1497
Exact Tensor Completion Using t-SVD http://dx.doi.org/10.1109/TSP.2016.2639466	<i>Z. Zhang and S. Aeron</i>	1511
Slepian Spatial-Spectral Concentration Problem on the Sphere: Analytical Formulation for Limited Colatitude-Longitude Spatial Region http://dx.doi.org/10.1109/TSP.2016.2646668	<i>A. P. Bates, Z. Khalid, and R. A. Kennedy</i>	1527
On Lower Bounds for Nonstandard Deterministic Estimation http://dx.doi.org/10.1109/TSP.2016.2645538	<i>N. Kbayer, J. Galy, E. Chaumette, F. Vincent, A. Renaux, and P. Larzabal</i>	1538
On Lossless Feedback Delay Networks http://dx.doi.org/10.1109/TSP.2016.2637323	<i>S. J. Schlecht and E. A. P. Habets</i>	1554
Recursive Least Squares for Censored Regression http://dx.doi.org/10.1109/TSP.2016.2646660	<i>Z. Liu and C. Li</i>	1565
Randomized Robust Subspace Recovery and Outlier Detection for High Dimensional Data Matrices http://dx.doi.org/10.1109/TSP.2016.2645515	<i>M. Rahmani and G. K. Atia</i>	1580
BIBO-Stable Recursive Functional Link Polynomial Filters http://dx.doi.org/10.1109/TSP.2016.2641395	<i>A. Carini and G. L. Sicuranza</i>	1595
Performance of Uniform and Sparse Non-Uniform Samplers in Presence of Modeling Errors: A Cramér-Rao Bound Based Study http://dx.doi.org/10.1109/TSP.2016.2637309	<i>A. Koochakzadeh and P. Pal</i>	1607
A Soft-Output MIMO Detector With Achievable Information Rate Based Partial Marginalization http://dx.doi.org/10.1109/TSP.2016.2641393	<i>S. Hu and F. Rusek</i>	1622

CORRECTIONS

Corrections to “Quantized Consensus by the ADMM: Probabilistic Versus Deterministic Quantizers” http://dx.doi.org/10.1109/TSP.2016.2641305	<i>S. Zhu and B. Chen</i>	1638
--	---------------------------	------



CALL FOR PAPERS

MMSP 2017 is the IEEE 19th International Workshop on Multimedia Signal Processing. The workshop is organized by the Multimedia Signal Processing Technical Committee of the IEEE Signal Processing Society. This year's event has a theme of '**Multimedia Processing for Healthcare and Assisted Living**'.

Recent advances in multimedia processing and communications have potential to significantly advance current healthcare and assisted living services by enabling remote health monitoring, remote diagnostics, increased patient privacy, robotic-assisted surgery, and home-based treatment. A huge diversity of multimedia processing techniques, ranging from image/audio sensing, compression and networking, denoising, feature extraction, security, distributed processing, depth image processing, cloud and social computing, visualization and multimedia big data analytics, can all find their applications in future healthcare and elderly-care. However, to fully realize this potential and embed multimedia solutions into day-to-day healthcare practice, many challenges need to be overcome that call for significant engineering innovation, which can only happen through close interdisciplinary effort. The MMSP-2017 Workshop will bring together experts from different fields, including signal processing, computer science, communications, medicine, rehabilitation, psychology, to exchange ideas on how multimedia research can support advancements of future healthcare and how best to facilitate interactions between multimedia researchers and healthcare and assisted living providers.

Papers are solicited in (but not limited to) the following areas, covering not only this year's workshop theme, but also the general scope of multimedia signal processing:

- Multimedia big data analytics
- Distributed multimedia for body networks
- Deep learning for health-specific event detection and classification
- Streaming, security and privacy for healthcare
- Sparsity-based and low-rank based sensing of human vital signs
- Multimedia for smart homes and elderly care
- Multimedia processing for tele-rehabilitation
- Computational imaging for healthcare applications
- Healthcare monitoring applications using wearable technologies
- Image/video/speech/audio coding and processing
- Multimedia networking
- Multimedia traffic, communications and heterogeneous interactions
- Multimedia quality assessment
- Internet of Things (IoT)-based multimedia systems and applications
- Multimedia hardware design
- Augmented, mixed and virtual reality

Important dates:

Proposals for Special Sessions and Tutorials: **April 1, 2017**
 Notification of Acceptance for Special Session and Tutorial Proposals: **April 15, 2017**
 Submission of Regular and Special Session Papers: **June 1, 2017**
 Notification of Acceptance for Regular and Special Session Papers: **July 15, 2017**
 Submission of Sketch and Demo Papers: **August 1, 2017**
 Camera Ready Deadline: **August 1, 2017**

General Co-chairs:

Vladan VELISAVLJEVIC, *University of Bedfordshire, UK*
 Vladimir STANKOVIC, *University of Strathclyde, UK*
 Zixiang XIONG, *Texas A&M University, TX, USA*

TPC Co-chairs:

Homer H. CHEN, *National Taiwan University, Taiwan*
 Frederic DUFAUX, *CNRS, France*
 Raouf HAMZAOU, *De Montfort University, UK*

Finance Co-chairs:

Ching-Te CHIU, *National Tsing Hua University, Taiwan*
 Lina STANKOVIC, *University of Strathclyde, UK*

Special Session & Panel Co-chairs:

Gene CHEUNG, *National Institute for Informatics, Japan*
 Tasos DAGIUKLAS, *London South Bank University, UK*

Keynote Co-chairs:

Nikolaos BOULGOURIS, *Brunel University London, UK*
 Yiannis ANDREOPOULOS, *University College London, UK*

Industry & Demo Chair:

Marta MRAK, *BBC R&D, UK*

Publication Chair:

Maria MARTINI, *Kingston University London, UK*
 Safak DOGAN, *Loughborough University London, UK*

Publicity Chair:

Ivana TOSIC, *Ricoh, CA, USA*

Asia & Australia Liaison:

Tao MEI, *Microsoft Research Beijing, China*

North & South America Liaison:

Jacob CHAKARESKE, *University of Alabama, AL, USA*

Local Arrangements Chair:

Nicholas GARDNER, *University of Bedfordshire, UK*





IEEE TRANSACTIONS ON

SIGNAL AND INFORMATION PROCESSING OVER NETWORKS



Now accepting paper submissions

The new *IEEE Transactions on Signal and Information Processing over Networks* publishes high-quality papers that extend the classical notions of processing of signals defined over vector spaces (e.g. time and space) to processing of signals and information (data) defined over networks, potentially dynamically varying. In signal processing over networks, the topology of the network may define structural relationships in the data, or may constrain processing of the data. Topics of interest include, but are not limited to the following:

Adaptation, Detection, Estimation, and Learning

- Distributed detection and estimation
- Distributed adaptation over networks
- Distributed learning over networks
- Distributed target tracking
- Bayesian learning; Bayesian signal processing
- Sequential learning over networks
- Decision making over networks
- Distributed dictionary learning
- Distributed game theoretic strategies
- Distributed information processing
- Graphical and kernel methods
- Consensus over network systems
- Optimization over network systems

Communications, Networking, and Sensing

- Distributed monitoring and sensing
- Signal processing for distributed communications and networking
- Signal processing for cooperative networking
- Signal processing for network security
- Optimal network signal processing and resource allocation

Modeling and Analysis

- Performance and bounds of methods
- Robustness and vulnerability
- Network modeling and identification

Modeling and Analysis (cont.)

- Simulations of networked information processing systems
- Social learning
- Bio-inspired network signal processing
- Epidemics and diffusion in populations

Imaging and Media Applications

- Image and video processing over networks
- Media cloud computing and communication
- Multimedia streaming and transport
- Social media computing and networking
- Signal processing for cyber-physical systems
- Wireless/mobile multimedia

Data Analysis

- Processing, analysis, and visualization of big data
- Signal and information processing for crowd computing
- Signal and information processing for the Internet of Things
- Emergence of behavior

Emerging topics and applications

- Emerging topics
- Applications in life sciences, ecology, energy, social networks, economic networks, finance, social sciences, smart grids, wireless health, robotics, transportation, and other areas of science and engineering

Editor-in-Chief: Petar M. Djurić, Stony Brook University (USA)

To submit a paper, go to: <https://mc.manuscriptcentral.com/tsipn-ieee>



IEEE/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



FEBRUARY 2017

VOLUME 25

NUMBER 2

ITASFA

(ISSN 2329-9290)

REGULAR PAPERS

Direct-to-Reverberant Energy Ratio Estimation Using a First-Order Microphone http://dx.doi.org/10.1109/TASLP.2016.2601222	226
..... <i>H. Chen, T. D. Abhayapala, P. N. Samarasinghe, and W. Zhang</i>	
Multitask Learning of Context-Dependent Targets in Deep Neural Network Acoustic Models http://dx.doi.org/10.1109/TASLP.2016.2630305	238
..... <i>P. Bell, P. Swietojanski, and S. Renals</i>	
Topic-Aware Deep Compositional Models for Sentence Classification http://dx.doi.org/10.1109/TASLP.2016.2632521	248
..... <i>R. Zhao and K. Mao</i>	
On-the-Fly Audio Source Separation—A Novel User-Friendly Framework http://dx.doi.org/10.1109/TASLP.2016.2632528	261
..... <i>D. E. Badawy, N. Q. K. Duong, and A. Ozerov</i>	



Online Estimation of Multiple Harmonic Signals http://dx.doi.org/10.1109/TASLP.2016.2634118	<i>F. Elvander, J. Swärd, and A. Jakobsson</i>	273
Weakly Supervised Learning of Hidden Markov Models for Spoken Language Acquisition http://dx.doi.org/10.1109/TASLP.2016.2633812	<i>V. Renkens and H. V. Hamme</i>	285
Acoustic Reflector Localization: Novel Image Source Reversion and Direct Localization Methods http://dx.doi.org/10.1109/TASLP.2016.2633802	<i>L. Remaggi, P. J. B. Jackson, P. Coleman, and W. Wang</i>	296
Estimating the Direct-to-Reverberant Energy Ratio Using a Spherical Harmonics-Based Spatial Correlation Model http://dx.doi.org/10.1109/TASLP.2016.2633811	<i>P. N. Samarasinghe, T. D. Abhayapala, and H. Chen</i>	310
Combined LCMV-TRINICON Beamforming for Separating Multiple Speech Sources in Noisy and Reverberant Environments http://dx.doi.org/10.1109/TASLP.2016.2633806	<i>S. Markovich-Golan, S. Gannot, and W. Kellermann</i>	320
Gain Scheduling of Auxiliary Noise and Variable Step-Size for Online Acoustic Feedback Cancellation in Narrow-Band Active Noise Control Systems http://dx.doi.org/10.1109/TASLP.2016.2635029	<i>S. Ahmed and M. T. Akhtar</i>	333
Estimating the Structural Segmentation of Popular Music Pieces Under Regularity Constraints http://dx.doi.org/10.1109/TASLP.2016.2635031	<i>G. Sargent, F. Bimbot, and E. Vincent</i>	344
An Investigation of Delayless Subband Adaptive Filtering for Multi-Input Multi-Output Active Noise Control Applications http://dx.doi.org/10.1109/TASLP.2016.2637298	<i>J. Cheer and S. Daley</i>	359
Feedback Delay Networks: Echo Density and Mixing Time http://dx.doi.org/10.1109/TASLP.2016.2635027	<i>S. J. Schlecht and E. A. P. Habets</i>	374
An Instrumental Quality Measure for Artificially Bandwidth-Extended Speech Signals http://dx.doi.org/10.1109/TASLP.2016.2635022	<i>J. Abel, M. Kaniewska, C. Guillaumé, W. Tirry, and T. Fingscheidt</i>	384
An Analysis of Adaptive Recursive Smoothing with Applications to Noise PSD Estimation http://dx.doi.org/10.1109/TASLP.2016.2633803	<i>R. Rehr and T. Gerkman</i>	397
Multimodal Crowdsourcing for Transcribing Handwritten Documents http://dx.doi.org/10.1109/TASLP.2016.2634123	<i>E. Granell and C.-D. Martínez-Hinarejos</i>	409
A New Strategy for Online Secondary-Path Modeling of Narrowband Active Noise Control http://dx.doi.org/10.1109/TASLP.2016.2633799	<i>Y. Ma and Y. Xiao</i>	420
GPU-Based Dynamic Wave Field Synthesis Using Fractional Delay Filters and Room Compensation http://dx.doi.org/10.1109/TASLP.2016.2631338	<i>J. A. Belloch, A. Gonzalez, E. S. Quintana-Ortí, M. Ferrer, and V. Välimäki</i>	435

EDICS—Editor’s Information Classification Scheme http://dx.doi.org/10.1109/TASLP.2016.2651856		448
Information for Authors http://dx.doi.org/10.1109/TASLP.2016.2651858		450

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



FEBRUARY 2017

VOLUME 26

NUMBER 2

IIPRE4

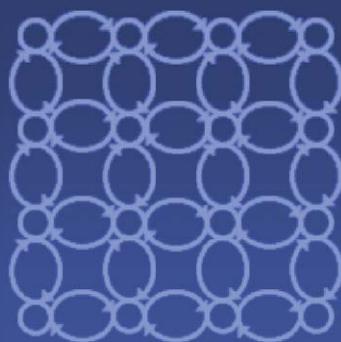
(ISSN 1057-7149)

PAPERS

Random Walk Graph Laplacian-Based Smoothness Prior for Soft Decoding of JPEG Images http://dx.doi.org/10.1109/TIP.2016.2627807 ..	509
..... X. Liu, G. Cheung, X. Wu, and D. Zhao	
Depth-Guided Disocclusion Inpainting of Synthesized RGB-D Images http://dx.doi.org/10.1109/TIP.2016.2619263	525
..... P. Buysens, O. L. Meur, M. Daisy, D. Tschumperlé, and O. Lézoray	
A Parallel Proximal Algorithm for Anisotropic Total Variation Minimization http://dx.doi.org/10.1109/TIP.2016.2629449 ...	539
..... U. S. Kamilov	
Texture Reconstruction Guided by a High-Resolution Patch http://dx.doi.org/10.1109/TIP.2016.2627812	549
..... M. El Gheche, J.-F. Aujol, Y. Berthoumieu, and C.-A. Deledalle	
The Shape Interaction Matrix-Based Affine Invariant Mismatch Removal for Partial-Duplicate Image Search http://dx.doi.org/10.1109/TIP.2016.2629452	561
..... Y. Lin, Z. Lin, and H. Zha	
Context Tree-Based Image Contour Coding Using a Geometric Prior http://dx.doi.org/10.1109/TIP.2016.2627813	574
..... A. Zheng, G. Cheung, and D. Florencio	
Planelets-A Piecewise Linear Fractional Model for Preserving Scene Geometry in Intra-Coding of Indoor Depth Images http://dx.doi.org/10.1109/TIP.2016.2621672	590
..... V. Kiani, A. Harati, and A. Vahedian	
Fast Physically Correct Refocusing for Sparse Light Fields Using Block-Based Multi-Rate View Interpolation http://dx.doi.org/10.1109/TIP.2016.2631904	603
..... C.-T. Huang, Y.-W. Wang, L.-R. Huang, J. Chin, and L.-G. Chen	
A Context-Driven Extractive Framework for Generating Realistic Image Descriptions http://dx.doi.org/10.1109/TIP.2016.2628585	619
..... A. Tariq and H. Foroosh	
A Joint Compression Scheme of Video Feature Descriptors and Visual Content http://dx.doi.org/10.1109/TIP.2016.2629447	633
..... X. Zhang, S. Ma, S. Wang, X. Zhang, H. Sun, and W. Gao	
Flexible Multi-View Dimensionality Co-Reduction http://dx.doi.org/10.1109/TIP.2016.2627806	648
..... C. Zhang, H. Fu, Q. Hu, P. Zhu, and X. Cao	
Robust Transfer Metric Learning for Image Classification http://dx.doi.org/10.1109/TIP.2016.2631887	660
..... Z. Ding and Y. Fu	
Stereo Object Proposals http://dx.doi.org/10.1109/TIP.2016.2627819	671
..... S. Huang, W. Wang, S. He, and R. W. H. Lau	
A Non-Greedy Algorithm for L1-Norm LDA http://dx.doi.org/10.1109/TIP.2016.2621667	684
..... Y. Liu, Q. Gao, S. Miao, X. Gao, F. Nie, and Y. Li	
Affine Covariant Features for Fisheye Distortion Local Modeling http://dx.doi.org/10.1109/TIP.2016.2627816	696
..... A. Furnari, G. M. Farinella, A. R. Bruna, and S. Battiato	



Structure-From-Motion in Spherical Video Using the von Mises-Fisher Distribution http://dx.doi.org/10.1109/TIP.2016.2621662	711
..... <i>H. Guan and W. A. P. Smith</i>	
Moving Object Detection Using Tensor-Based Low-Rank and Saliency Fused-Sparse Decomposition http://dx.doi.org/10.1109/TIP.2016.2627803	724
..... <i>W. Hu, Y. Yang, W. Zhang, and Y. Xie</i>	
Sparse Canonical Temporal Alignment With Deep Tensor Decomposition for Action Recognition http://dx.doi.org/10.1109/TIP.2016.2621664	738
..... <i>C. Jia, M. Shao, and Y. Fu</i>	
Multi-Modal Dictionary Learning for Image Separation With Application in Art Investigation http://dx.doi.org/10.1109/TIP.2016.2623484 ..	751
..... <i>N. Deligiannis, J. F. C. Mota, B. Cornelis, M. R. D. Rodrigues, and I. Daubechies</i>	
Sparse Representation-Based Multiple Frame Video Super-Resolution http://dx.doi.org/10.1109/TIP.2016.2631339	765
..... <i>Q. Dai, S. Yoo, A. Kappeler, and A. K. Katsaggelos</i>	
Sparsity-Based Image Error Concealment via Adaptive Dual Dictionary Learning and Regularization http://dx.doi.org/10.1109/TIP.2016.2623481	782
..... <i>X. Liu, D. Zhai, J. Zhou, S. Wang, D. Zhao, and H. Gao</i>	
Adaptive Cascade Regression Model for Robust Face Alignment http://dx.doi.org/10.1109/TIP.2016.2633939	797
..... <i>Q. Liu, J. Deng, J. Yang, G. Liu, and D. Tao</i>	
Locally Supervised Deep Hybrid Model for Scene Recognition http://dx.doi.org/10.1109/TIP.2016.2629443	808
..... <i>S. Guo, W. Huang, L. Wang, and Y. Qiao</i>	
Automatic Refinement Strategies for Manual Initialization of Object Trackers http://dx.doi.org/10.1109/TIP.2016.2633874	821
..... <i>H. Zhu and F. Porikli</i>	
Exploiting Depth From Single Monocular Images for Object Detection and Semantic Segmentation http://dx.doi.org/10.1109/TIP.2016.2621673	836
..... <i>Y. Cao, C. Shen, and H. T. Shen</i>	
Tree Structure Sparsity Pattern Guided Convex Optimization for Compressive Sensing of Large-Scale Images http://dx.doi.org/10.1109/TIP.2016.2633872	847
..... <i>W.-J. Liang, G.-X. Lin, and C.-S. Lu</i>	
Turning Diffusion-Based Image Colorization Into Efficient Color Compression http://dx.doi.org/10.1109/TIP.2016.2627800	860
..... <i>P. Peter, L. Kaufhold, and J. Weickert</i>	
Automatic Design of High-Sensitivity Color Filter Arrays With Panchromatic Pixels http://dx.doi.org/10.1109/TIP.2016.2633869	870
..... <i>J. Li, C. Bai, Z. Lin, and J. Yu</i>	
Modality-Invariant Image Classification Based on Modality Uniqueness and Dictionary Learning http://dx.doi.org/10.1109/TIP.2016.2635444	884
..... <i>S. Kim, R. Cai, K. Park, S. Kim, and K. Sohn</i>	
Guided Wavelet Shrinkage for Edge-Aware Smoothing http://dx.doi.org/10.1109/TIP.2016.2633941	900
..... <i>G. Deng</i>	
Image Quality Assessment Based on Local Linear Information and Distortion-Specific Compensation http://dx.doi.org/10.1109/TIP.2016.2639451	915
..... <i>H. Wang, J. Fu, W. Lin, S. Hu, C.-C. Jay Kuo, and L. Zuo</i>	
ORGM: Occlusion Relational Graphical Model for Human Pose Estimation http://dx.doi.org/10.1109/TIP.2016.2639441	927
..... <i>L. Fu, J. Zhang, and K. Huang</i>	
Generalizing Mumford-Shah Model for Multiphase Piecewise Smooth Image Segmentation http://dx.doi.org/10.1109/TIP.2016.2636450 ...	942
..... <i>Y. Gu, W. Xiong, L.-L. Wang, and J. Cheng</i>	
Adaptive Multispectral Demosaicking Based on Frequency-Domain Analysis of Spectral Correlation http://dx.doi.org/10.1109/TIP.2016.2634120	953
..... <i>S. P. Jaiswal, L. Fang, V. Jakhetiya, J. Pang, K. Mueller, and O. C. Au</i>	
Manifold Regularized Experimental Design for Active Learning http://dx.doi.org/10.1109/TIP.2016.2635440	969
..... <i>L. Zhang, H. P. H. Shum, and L. Shao</i>	
LIME: Low-Light Image Enhancement via Illumination Map Estimation http://dx.doi.org/10.1109/TIP.2016.2639450	982
..... <i>X. Guo, Y. Li, and H. Ling</i>	
Pairwise Operator Learning for Patch-Based Single-Image Super-Resolution http://dx.doi.org/10.1109/TIP.2016.2639440	994
..... <i>Y. Tang and L. Shao</i>	
Waterloo Exploration Database: New Challenges for Image Quality Assessment Models http://dx.doi.org/10.1109/TIP.2016.2631888	1004
..... <i>K. Ma, Z. Duanmu, Q. Wu, Z. Wang, H. Yong, H. Li, and L. Zhang</i>	
Noise Level Estimation for Natural Images Based on Scale-Invariant Kurtosis and Piecewise Stationarity http://dx.doi.org/10.1109/TIP.2016.2639447	1017
..... <i>L. Dong, J. Zhou, and Y. Y. Tang</i>	
A New Intrinsic-Lighting Color Space for Daytime Outdoor Images http://dx.doi.org/10.1109/TIP.2016.2642788	1031
..... <i>Z. Han, J. Tian, L. Qu, and Y. Tang</i>	
Fast and Exact Newton and Bidirectional Fitting of Active Appearance Models http://dx.doi.org/10.1109/TIP.2016.2642828	1040
..... <i>J. Kossaiji, G. Tzimiropoulos, and M. Pantic</i>	
<hr/>	
EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIP.2016.2649413	1054
Information for Authors http://dx.doi.org/10.1109/TIP.2016.2649412	1055



MLSP2017

IEEE International Workshop on
MACHINE LEARNING FOR SIGNAL PROCESSING

SEPTEMBER 25-28, 2017

ROPPONGI, TOKYO, JAPAN

MLSP2017.CONWIZ.DK

Organizing committee

General Chairs	Tomoko Matsui The Institute of Statistical Mathematics, Japan Jen-Tzung Chien National Chiao Tung University, Taiwan
Program Chairs	Naonori Ueda NTT Communication Science Laboratories, Japan Shinji Watanabe Mitsubishi Electric Research Laboratories, Japan
Finance Chair	Motoi Okamoto The Institute of Statistical Mathematics, Japan
Publicity Chair	Marc van Hulle <i>KU Leuven, Belgium</i>
Publication Chair	Jan Larsen <i>Technical University of Denmark, Denmark</i>
Program Advisory Committee	Shigeru Katagiri Doshisha University, Japan Tülay Adalı University of Maryland, Baltimore County, USA Vince Calhoun University of New Mexico, USA Raviv Raich Oregon State University, USA



Copyright by Moyan Brenn, CC-BY-2.0

CALL FOR PAPERS

The 27th MLSP workshop in the series of workshops organized by the IEEE Signal Processing Society MLSP Technical Committee will present the most recent and exciting advances in machine learning for signal processing through keynote talks, tutorials, as well as special and regular single-track sessions. Prospective authors are invited to submit papers on relevant algorithms and applications including, but not limited to:

- Bayesian learning and signal processing
- Cognitive information processing
- Deep learning techniques
- Dictionary learning
- Graphical and kernel methods
- Independent component analysis
- Information-theoretic learning
- Learning theory and algorithms
- Pattern recognition and classification
- Bounds on performance
- Subspace and manifold learning
- Sequential learning and decision methods
- Source separation
- Tensor and structured matrix methods
- Machine learning from big data
- Scalable learning algorithms
- Applications including: speech, audio & music, image & video, biomedical signals & images, communications, bioinformatics, biometrics, systems biology, computational intelligence, genomic signals & sequences, social networks, games, smart grid, security & privacy

Important Dates and Deadlines

Special session proposals	April 3, 2017
Paper submissions	May 28, 2017
Decision notifications	July 5, 2017
Camera-ready papers and advance author registration	August 7, 2017

Sponsored by



Supported by



Research Organization of Information and Systems

The Institute of Statistical Mathematics

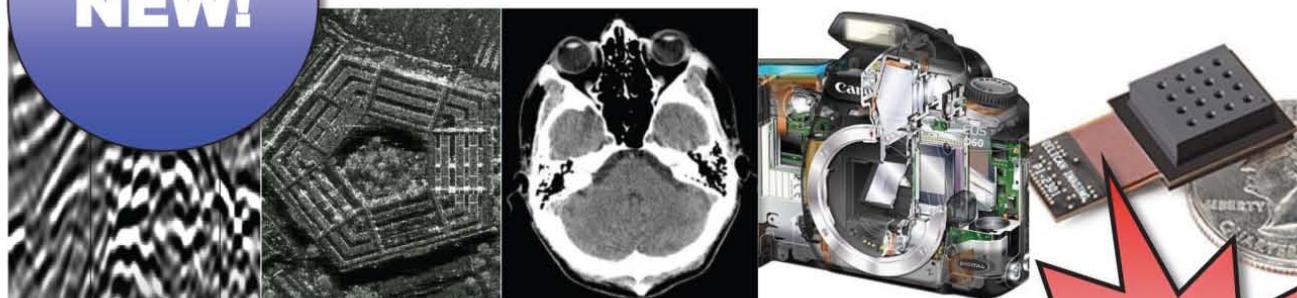
MLSP 2017 seeks proposals for **Special Sessions** that will address research in emerging or interdisciplinary areas of particular interest, not covered already by traditional MLSP sessions.

The **MLSP Best Student Paper Award** will be granted to the best paper for which a student is the principal author and presenter.

Prospective authors are invited to submit a double column paper of up to six pages using the electronic submission procedure at <http://mlsp2017.conwiz.dk>. Accepted papers will be published on a password-protected website that will be available during the workshop. The presented papers will be published in and indexed by IEEE Xplore.



IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING



Editor-in-Chief

W. Clem Karl
Boston University

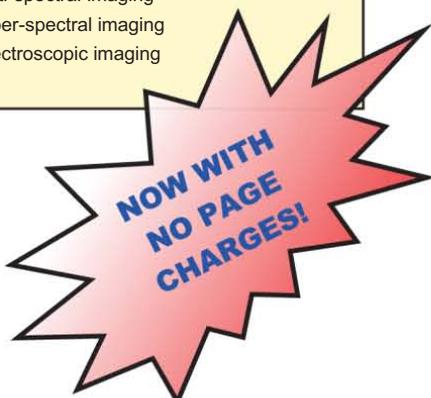
Technical Committee

Charles Bouman
Eric Miller
Peter Corcoran
Jong Chul Ye
Dave Brady
William Freeman

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:

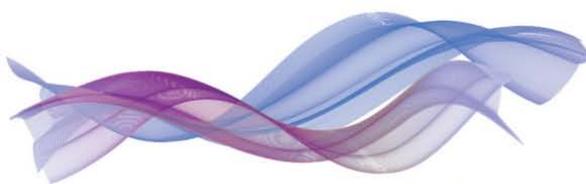


<p>Computational Imaging Methods and Models</p> <ul style="list-style-type: none"> • Coded image sensing • Compressed sensing • Sparse and low-rank models • Learning-based models, dictionary methods • Graphical image models • Perceptual models <p>Computational Image Formation</p> <ul style="list-style-type: 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@uoit.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

FEBRUARY 2017

VOLUME 12

NUMBER 2

ITIFA6

(ISSN 1556-6013)

REGULAR PAPERS

Secrecy Analysis of MIMO Wiretap Channels With Low-Complexity Receivers Under Imperfect Channel Estimation http://dx.doi.org/10.1109/TIFS.2016.2604490 ... <i>F. S. Al-Qahtani, Y. Huang, S. Hessian, R. M. Radaydeh, C. Zhong, and H. M. Alnuweiri</i>	257
Dynamic Privacy Pricing: A Multi-Armed Bandit Approach With Time-Variant Rewards http://dx.doi.org/10.1109/TIFS.2016.2611487 <i>L. Xu, C. Jiang, Y. Qian, Y. Zhao, J. Li, and Y. Ren</i>	271
We Can Track You if You Take the Metro: Tracking Metro Riders Using Accelerometers on Smartphones http://dx.doi.org/10.1109/TIFS.2016.2611489 <i>J. Hua, Z. Shen, and S. Zhong</i>	286
On Simultaneously Generating Multiple Keys in a Joint Source-Channel Model http://dx.doi.org/10.1109/TIFS.2016.2612172 <i>W. Tu, M. Goldenbaum, L. Lai, and H. V. Poor</i>	298
Exploration of Benes Network in Cryptographic Processors: A Random Infection Countermeasure for Block Ciphers Against Fault Attacks http://dx.doi.org/10.1109/TIFS.2016.2612638 <i>B. Wang, L. Liu, C. Deng, M. Zhu, S. Yin, Z. Zhou, and S. Wei</i>	309
Multi-Party Secret Key Agreement Over State-Dependent Wireless Broadcast Channels http://dx.doi.org/10.1109/TIFS.2016.2612649 <i>M. Jafari Siavoshani, S. Mishra, C. Fragouli, and S. N. Diggavi</i>	323
COTD: Reference-Free Hardware Trojan Detection and Recovery Based on Controllability and Observability in Gate-Level Netlist http://dx.doi.org/10.1109/TIFS.2016.2613842 <i>H. Salmani</i>	338
On Channel-Aware Secure HARQ-IR http://dx.doi.org/10.1109/TIFS.2016.2613846 <i>J. Choi</i>	351
Static and Dynamic Obfuscations of Scan Data Against Scan-Based Side-Channel Attacks http://dx.doi.org/10.1109/TIFS.2016.2613847 ... <i>A. Cui, Y. Luo, and C.-H. Chang</i>	363
Image Alignment-Based Multi-Region Matching for Object-Level Tampering Detection http://dx.doi.org/10.1109/TIFS.2016.2615272 <i>C.-M. Pun, C. Yan, and X.-C. Yuan</i>	377
Sensor Pattern Noise Estimation Based on Improved Locally Adaptive DCT Filtering and Weighted Averaging for Source Camera Identification and Verification http://dx.doi.org/10.1109/TIFS.2016.2620280 <i>A. Lawgaly and F. Khelifi</i>	392
Eliminating the Hardware-Software Boundary: A Proof-Carrying Approach for Trust Evaluation on Computer Systems http://dx.doi.org/10.1109/TIFS.2016.2621999 <i>X. Guo, R. G. Dutta, and Y. Jin</i>	405
Preserving the Location Privacy of Secondary Users in Cooperative Spectrum Sensing http://dx.doi.org/10.1109/TIFS.2016.2622000 <i>M. Grissa, A. A. Yavuz, and B. Hamdaoui</i>	418
Detection of Double Compressed AMR Audio Using Stacked Autoencoder http://dx.doi.org/10.1109/TIFS.2016.2622012 <i>D. Luo, R. Yang, B. Li, and J. Huang</i>	432



Using the Spanning Tree of a Criminal Network for Identifying Its Leaders http://dx.doi.org/10.1109/TIFS.2016.2622226	445
..... <i>K. Taha and P. D. Yoo</i>	
Efficient Hierarchical Identity-Based Signature With Batch Verification for Automatic Dependent Surveillance- Broadcast System http://dx.doi.org/10.1109/TIFS.2016.2622682	454
..... <i>D. He, N. Kumar, K.-K. R. Choo, and W. Wu</i>	
A Steganalytic Approach to Detect Motion Vector Modification Using Near-Perfect Estimation for Local Optimality http://dx.doi.org/10.1109/TIFS.2016.2623587	465
..... <i>H. Zhang, Y. Cao, and X. Zhao</i>	
Optimized 3D Lighting Environment Estimation for Image Forgery Detection http://dx.doi.org/10.1109/TIFS.20162.2623589	479
..... <i>B. Peng, W. Wang, J. Dong, and T. Tan</i>	

EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIFS.2017.2652410	495
Information for Authors http://dx.doi.org/10.1109/TIFS.2017.2652411	496

ANNOUNCEMENTS

Call for Papers-IEEE Journal of Selected Topics in Signal Processing Special Issue on End-to-End Speech and Language Processing http://dx.doi.org/10.1109/TIFS.2017.2652412	498
---	-----

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY



www.signalprocessingsociety.org

MARCH 2017

VOLUME 12

NUMBER 3

ITIFA6

(ISSN 1556-6013)

REGULAR PAPERS

Secrecy Throughput Maximization for MISO Multi-Eavesdropper Wiretap Channels http://dx.doi.org/10.1109/TIFS.2016.2620279	505
..... <i>W. Wang, K. C. Teh, and K. H. Li</i>	
On the Role of Artificial Noise in Training and Data Transmission for Secret Communications http://dx.doi.org/10.1109/TIFS.2016.2620281	516
..... <i>T.-Y. Liu, S.-C. Lin, and Y.-W. P. Hong</i>	
Secure and Efficient Cloud Data Deduplication With Randomized Tag http://dx.doi.org/10.1109/TIFS.2016.2622013	532
..... <i>T. Jiang, X. Chen, Q. Wu, J. Ma, W. Susilo, and W. Lou</i>	
Crime Scene Reconstruction: Online Gold Farming Network Analysis http://dx.doi.org/10.1109/TIFS.2016.2623586	544
..... <i>H. Kwon, A. Mohaisen, J. Woo, Y. Kim, E. Lee, and H. K. Kim</i>	
On Resilience and Connectivity of Secure Wireless Sensor Networks Under Node Capture Attacks http://dx.doi.org/10.1109/TIFS.2016.2613841	557
..... <i>J. Zhao</i>	
On the Secrecy Capacity Region of the Two-User Symmetric Z Interference Channel With Unidirectional Transmitter Cooperation http://dx.doi.org/10.1109/TIFS.2016.2622007	572
..... <i>P. Mohapatra, C. R. Murthy, and J. Lee</i>	
Smart Detection of Line-Search Oracle Attacks http://dx.doi.org/10.1109/TIFS.2016.2624280	588
..... <i>B. Tondi, P. Comesaña-Alfaro, F. Pérez-González, and M. Barni</i>	
FACT: A Framework for Authentication in Cloud-Based IP Traceback http://dx.doi.org/10.1109/TIFS.2016.2624741	604
..... <i>L. Cheng, D. M. Divakaran, A. Wooi Kiak Ang, W. Y. Lim, and V. L. L. Thing</i>	
Secrecy Capacity Analysis of Artificial Noisy MIMO Channels-An Approach Based on Ordered Eigenvalues of Wishart Matrices http://dx.doi.org/10.1109/TIFS.2016.2627219	617
..... <i>Y. Liu, H.-H. Chen, and L. Wang</i>	
ALDOCX: Detection of Unknown Malicious Microsoft Office Documents Using Designated Active Learning Methods Based on New Structural Feature Extraction Methodology http://dx.doi.org/10.1109/TIFS.2016.2631905	631
..... <i>N. Nissim, A. Cohen, and Y. Elovici</i>	
Identifying User-Input Privacy in Mobile Applications at a Large Scale http://dx.doi.org/10.1109/TIFS.2016.2631949	647
..... <i>Y. Nan, Z. Yang, M. Yang, S. Zhou, Y. Zhang, G. Gu, X. Wang, and L. Sun</i>	
Light-Weight and Robust Security-Aware D2D-Assist Data Transmission Protocol for Mobile-Health Systems http://dx.doi.org/10.1109/TIFS.2016.2631950	662
..... <i>A. Zhang, L. Wang, X. Ye, and X. Lin</i>	
Efficient Public Verification of Data Integrity for Cloud Storage Systems from Indistinguishability Obfuscation http://dx.doi.org/10.1109/TIFS.2016.2631951	676
..... <i>Y. Zhang, C. Xu, X. Liang, H. Li, Y. Mu, and X. Zhang</i>	
Group Sparsity Tensor Factorization for Re-Identification of Open Mobility Traces http://dx.doi.org/10.1109/TIFS.2016.2631952	689
..... <i>T. Murakami, A. Kanemura, and H. Hino</i>	
A Novel Online Signature Verification System Based on GMM Features in a DTW Framework http://dx.doi.org/10.1109/TIFS.2016.2632063	705
..... <i>A. Sharma and S. Sundaram</i>	
Gender Estimation Based on Smile-Dynamics http://dx.doi.org/10.1109/TIFS.2016.2632070	719
..... <i>A. Dantcheva and F. Brémond</i>	
Detecting Anomalous Behavior in VoIP Systems: A Discrete Event System Modeling http://dx.doi.org/10.1109/TIFS.2016.2632071	730
..... <i>D. Golait and N. Hubballi</i>	
Finite State Markov Wiretap Channel With Delayed Feedback http://dx.doi.org/10.1109/TIFS.2016.2636085	746
..... <i>B. Dai, Z. Ma, and Y. Luo</i>	



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/>

FEBRUARY 2017

VOLUME 19

NUMBER 2

ITMUF8

(ISSN 1520-9210)

PAPERS

Low-Complexity Processing in Devices and Sensors

Fast, Compact, and Discriminative: Evaluation of Binary Descriptors for Mobile Applications
<http://dx.doi.org/10.1109/TMM.2016.2615521> *S. Madeo and M. Bober* 221

Compression and Coding

Wavelet-Based L_∞ Semi-regular Mesh Coding <http://dx.doi.org/10.1109/TMM.2016.2614483>
..... *R. Florea, A. Munteanu, S.-P. Lu, and P. Schelkens* 236

Watermarking, Encryption, and Data Hiding

Toward Encrypted Cloud Media Center With Secure Deduplication <http://dx.doi.org/10.1109/TMM.2016.2612760>
..... *Y. Zheng, X. Yuan, X. Wang, J. Jiang, C. Wang, and X. Gui* 251

Image/Video/Graphics Analysis and Synthesis

Fast and Adaptive 3D Reconstruction With Extensively High Completeness <http://dx.doi.org/10.1109/TMM.2016.2612761>
..... *P. Wu, Y. Liu, M. Ye, J. Li, and S. Du* 266

A Novel Transient Wrinkle Detection Algorithm and Its Application for Expression Synthesis
<http://dx.doi.org/10.1109/TMM.2016.2614429> *W. Xie, L. Shen, and J. Jiang* 279

Sparse Signal Processing

Color-Guided Depth Recovery via Joint Local Structural and Nonlocal Low-Rank Regularization
<http://dx.doi.org/10.1109/TMM.2016.2613824> *W. Dong, G. Shi, X. Li, K. Peng, J. Wu, and Z. Guo* 293

Retrieval Compensated Group Structured Sparsity for Image Super-Resolution <http://dx.doi.org/10.1109/TMM.2016.2614427>
..... *J. Liu, W. Yang, X. Zhang, and Z. Guo* 302



Integration of Synthetic and Natural Audio/Video

A Joint Deep Boltzmann Machine (jDBM) Model for Person Identification Using Mobile Phone Data
<http://dx.doi.org/10.1109/TMM.2016.2615524> M. R. Alam, M. Bennamoun, R. Togneri, and F. Sohel 317

Subjective and Objective Quality Assessment, and User Experience

Automatic Mesh Animation Preview With User Voting-Based Refinement <http://dx.doi.org/10.1109/TMM.2016.2612124> Z. Zhou, N. Jiang, K. Chen, and J. Zhang 327

Implicit Analysis of Perceptual Multimedia Experience Based on Physiological Response: A Review
<http://dx.doi.org/10.1109/TMM.2016.2614880> S.-E. Moon and J.-S. Lee 340

Content Description and Annotation

Multimodal Video-to-Near-Scene Annotation <http://dx.doi.org/10.1109/TMM.2016.2614426> C.-L. Chou, H.-T. Chen, and S.-Y. Lee 354

Knowledge and Semantics Modeling for Multimedia Databases

Joint Image-Text News Topic Detection and Tracking by Multimodal Topic And-Or Graph <http://dx.doi.org/10.1109/TMM.2016.2616279> ..
 W. Li, J. Joo, H. Qi, and S.-C. Zhu 367

Multimedia Search and Retrieval

Hashing With Pairwise Correlation Learning and Reconstruction <http://dx.doi.org/10.1109/TMM.2016.2614858> X.-J. Mao, Y.-B. Yang, and N. Li 382

Social and Web Multimedia

Weakly Supervised Learning of Deformable Part-Based Models for Object Detection via Region Proposals
<http://dx.doi.org/10.1109/TMM.2016.2614862> Y. Tang, X. Wang, E. Dellandrea, and L. Chen 393

Cyber-Physical-Social Spaces for Multimedia Applications and Smart City

Integration of Diverse Data Sources for Spatial PM2.5 Data Interpolation <http://dx.doi.org/10.1109/TMM.2016.2613639> M. Tang, X. Wu, P. Agrawal, S. Pongpaichet, and R. Jain 408

Multimedia for Personal Applications (Mobile, Wearables, Interactive)

Visualizing Video Sounds With Sound Word Animation to Enrich User Experience <http://dx.doi.org/10.1109/TMM.2016.2613641> F. Wang, H. Nagano, K. Kashino, and T. Igarashi 418

User-Centric Social Multimedia Computing and Location-Based Multimedia Services

Modeling Restaurant Context for Food Recognition <http://dx.doi.org/10.1109/TMM.2016.2614861> L. Herranz, S. Jiang, and R. Xu 430

Information for Authors <http://dx.doi.org/10.1109/TMM.2016.2648063> 441

ANNOUNCEMENTS

Introducing IEEE Collabratec <http://dx.doi.org/10.1109/TMM.2016.2649359> 443

IEEE Member Get-A-Member (MGM) Program <http://dx.doi.org/10.1109/TMM.2016.2649361> 444

2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2017)

October 15–18, 2017



The 2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2017) will be held at the Mohonk Mountain House in New Paltz, New York, and is supported by the Audio and Acoustic Signal Processing technical committee of the IEEE Signal Processing Society. The objective of this workshop is to provide an informal environment for the discussion of problems in audio, acoustics and signal processing techniques leading to novel solutions. Technical sessions will be scheduled throughout the day. Afternoons will be left free for informal meetings among workshop participants. Papers describing original research and new concepts are solicited for technical sessions on, but not limited to, the following topics:

Acoustic Signal Processing

- Source separation: single- and multi-microphone techniques
- Acoustic source localization and tracking
- Signal enhancement: dereverberation, noise reduction, echo reduction
- Microphone and loudspeaker array processing
- Acoustic sensor networks: distributed algorithms, synchronization
- Acoustic scene analysis: event detection and classification
- Room acoustics: analysis, modeling and simulation

Audio and Music Signal Processing

- Content-based music retrieval: fingerprinting, matching, cover song retrieval
- Musical signal analysis: segmentation, classification, transcription
- Music signal synthesis: waveforms, instrument models, singing
- Music separation: direct-ambient decomposition, vocal and instruments
- Audio effects: artificial reverberation, amplifier modeling
- Upmixing and downmixing

Audio and Speech Coding

- Waveform and parametric coding
- Spatial audio coding
- Sparse representations
- Low-delay audio and speech coding
- Digital rights

Hearing and Perception

- Hearing aids
- Computational auditory scene analysis
- Auditory perception and spatial hearing
- Speech and audio quality assessment
- Speech intelligibility measures and prediction

Important Dates

Submission of four-page paper
April 20, 2017

Notification of acceptance
June 27, 2017

Early registration until
August 15, 2017

Workshop
October 15–18, 2017

www.waspaa.com

Mohonk Mountain House
New Paltz, New York, USA

Workshop Committee

General Chairs

Patrick A. Naylor
Imperial College London

Meinard Müller
International AudioLabs Erlangen

Technical Program Chairs

Gautham Mysore
Adobe Research

Mads Christensen
Aalborg University

Finance Chair

Michael S. Brandstein
M.I.T. Lincoln Laboratory

Publications Chair

Toon van Waterschoot
KU Leuven

Registration Chair

Tiago H. Falk
INRS, Montréal

Industrial Liaison Chair

Tao Zhang
Starkey Hearing Technologies

Far East Liaison

Shoko Araki
NTT

Local Arrangements Chair

Youngjune Gwon
M.I.T. Lincoln Laboratory

Demonstrations Chair

Christine Evers
Imperial College London

Awards Chair

Sebastian Ewert
Queen Mary University of London



IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING


www.ieee.org/sp/index.html

DECEMBER 2016

VOLUME 10

NUMBER 8

IJSTGY

(ISSN 1932-4553)

ISSUE ON SIGNAL PROCESSING FOR EXPLOITING INTERFERENCE TOWARD ENERGY EFFICIENT AND SECURE WIRELESS COMMUNICATIONS

EDITORIAL

Introduction to the Issue on Signal Processing for Exploiting Interference Toward Energy Efficient and Secure Wireless Communications <http://dx.doi.org/10.1109/IJSTSP.2016.2620703> *I. Krikidis, C. Masouros, G. Zheng, R. Zhang, and R. Schober* 1331

PAPERS

Phase-Only Zero Forcing for Secure Communication With Multiple Antennas <http://dx.doi.org/10.1109/IJSTSP.2016.2611483> *W. Zhao, S.-H. Lee, and A. Khisti* 1334

A Collaboration Incentive Exploiting the Primary-Secondary Systems' Cross Interference for PHY Security Enhancement <http://dx.doi.org/10.1109/IJSTSP.2016.2600514> *K. Tourki and M. O. Hasna* 1346

Sum Secrecy Rate Maximization for Full-Duplex Two-Way Relay Networks Using Alamouti-Based Rank-Two Beamforming <http://dx.doi.org/10.1109/IJSTSP.2016.2603970> *Q. Li, W.-K. Ma, and D. Han* 1359

Secrecy and Energy Efficiency in Massive MIMO Aided Heterogeneous C-RAN: A New Look at Interference <http://dx.doi.org/10.1109/IJSTSP.2016.2600520> *L. Wang, K.-K. Wong, M. ElKashlan, A. Nallanathan, and S. Lambotharan* 1375

Secrecy Capacity Scaling by Jamming-Aided Hierarchical Cooperation in Ad Hoc Networks <http://dx.doi.org/10.1109/IJSTSP.2016.2616842> ..
..... *M. G. Kang, Y.-b. Kim, J. H. Lee, and W. Choi* 1390

Regularized Channel Inversion for Simultaneous Confidential Broadcasting and Power Transfer: A Large System Analysis <http://dx.doi.org/10.1109/IJSTSP.2016.2608792> *B. He, N. Yang, S. Yan, and X. Zhou* 1404



Secrecy Rate Optimization for Secure Multicast Communications http://dx.doi.org/10.1109/JSTSP.2016.2600518	1417
..... <i>K. Cumanan, Z. Ding, M. Xu, and H. V. Poor</i>	
Improving Wireless Physical Layer Security via Exploiting Co-Channel Interference http://dx.doi.org/10.1109/JSTSP.2016.2600516	1433
..... <i>L. Li, A. P. Petropulu, Z. Chen, and J. Fang</i>	
Wireless Information Surveillance via Proactive Eavesdropping with Spoofing Relay http://dx.doi.org/10.1109/JSTSP.2016.2600519	1449
..... <i>Y. Zeng and R. Zhang</i>	
Energy Efficiency of Confidential Multi-Antenna Systems With Artificial Noise and Statistical CSI http://dx.doi.org/10.1109/JSTSP.2016.2607690	1462
..... <i>A. Zappone, P.-H. Lin, and E. Jorswieck</i>	
Directional Modulation Via Symbol-Level Precoding: A Way to Enhance Security http://dx.doi.org/10.1109/JSTSP.2016.2600521	1478
..... <i>A. Kalantari, M. Soltanalian, S. Maleki, S. Chatzinotas, and B. Ottersten</i>	
Secure Multiple Amplify-and-Forward Relaying With Cochannel Interference http://dx.doi.org/10.1109/JSTSP.2016.2607692	1494
..... <i>L. Fan, X. Lei, N. Yang, T. Q. Duong, and G. K. Karagiannidis</i>	
Multipair Two-Way Relay Network With Harvest-Then-Transmit Users: Resolving Pairwise Uplink-Downlink Coupling http://dx.doi.org/10.1109/JSTSP.2016.2612162	1506
..... <i>S. Wang, M. Xia, and Y.-C. Wu</i>	
On the Performance of mmWave Networks Aided by Wirelessly Powered Relays http://dx.doi.org/10.1109/JSTSP.2016.2610944	1522
..... <i>S. Biswas, S. Vuppala, and T. Ratnarajah</i>	
Accumulate and Jam: Towards Secure Communication via A Wireless-Powered Full-Duplex Jammer http://dx.doi.org/10.1109/JSTSP.2016.2600523	1538
..... <i>Y. Bi and H. Chen</i>	
Quadrature Amplitude Modulation Division for Multiuser MISO Broadcast Channels http://dx.doi.org/10.1109/JSTSP.2016.2607684	1551
..... <i>Z. Dong, Y.-Y. Zhang, J.-K. Zhang, and X.-C. Gao</i>	

CORRECTION

Corrections to “A Characterization of Deterministic Sampling Patterns for Low-Rank Matrix Completion” http://dx.doi.org/10.1109/JSTSP.2016.2618502	1567
..... <i>D. L. Pimentel-Alarcón, N. Boston, and R. D. Nowak</i>	

List of Reviewers http://dx.doi.org/10.1109/JSTSP.2016.2618498	1568
---	------

Information for Authors http://dx.doi.org/10.1109/JSTSP.2016.2622188	1575
---	------

2016 INDEX HTTP://DX.DOI.ORG/10.1109/JSTSP.2016.2628027	1577
--	------

CALL FOR PAPERS

IEEE Journal of Selected Topics in Signal Processing

Special Issue on *End-to-End Speech and Language Processing*

End-to-end (E2E) systems have achieved competitive results compared to conventional hybrid Hidden Markov-deep neural network model-based automatic speech recognition (ASR) systems. Such E2E systems are attractive because they do not require initial alignments between input acoustic features and output graphemes or words. Very deep convolutional networks and recurrent neural networks have also been very successful in ASR systems due to their added expressive power and better generalization. ASR is often not the end goal of real-world speech information processing systems. Instead, an important end goal is information retrieval, in particular keyword search (KWS), which involves retrieving speech documents containing a user-specified query from a large database. Conventional keyword search uses an ASR system as a front-end that converts the speech database into a finite-state transducer (FST) index containing a large number of likely word or sub-word sequences for each speech segment, along with associated confidence scores and time stamps. A user-specified text query is then composed with this FST index to find the putative locations of the keyword along with confidence scores. More recently, inspired by E2E approaches, ASR-free keyword search systems have been proposed with limited success. Machine learning methods have also been very successful in Question-Answering, parsing, language translation, analytics and deriving representations of morphological units, words or sentences. Challenges such as the Zero Resource Speech Challenge aim to construct systems that learn an end-to-end Spoken Dialog (SD) system, in an unknown language, from scratch, using only information available to a language learning infant (zero linguistic resources). The principal objective of the recently concluded IARPA Babel program was to develop a keyword search system that delivers high accuracy for any new language given very limited transcribed speech, noisy acoustic and channel conditions, and limited system build time of one to four weeks. This special issue will showcase the power of novel machine learning methods not only for ASR, but for keyword search and for the general processing of speech and language.

Topics of interest in the special issue include (but are not limited to):

- Novel end-to-end speech and language processing
- Deep learning based acoustic and word representations
- Query-by-example search
- Question answering systems
- Multilingual dialogue systems
- Multilingual representation learning
- Low and zero resource speech processing
- Deep learning based ASR-free keyword search
- Deep learning based media retrieval
- Kernel methods applied to speech and language processing
- Acoustic unit discovery
- Computational challenges for deep end-to-end systems
- Adaptation strategies for end to end systems
- Noise robustness for low resource speech recognition
- Spoken language processing: speech retrieval, speech to speech translation, extraction, and summarization
- Machine learning methods applied to morphological, syntactic, and pragmatic analysis
- Computational semantics: document analysis, topic segmentation, categorization, and modeling
- Named entity recognition, tagging, chunking, and parsing
- Sentiment analysis, opinion mining, and social media analytics
- Deep learning in human computer interaction

Prospective authors should follow the instructions given on the IEEE JSTSP webpages: <https://signalprocessingsociety.org/publications-resources/ieee-journal-selected-topics-signal-processing>, and submit their manuscript with the web submission system at: <https://mc.manuscriptcentral.com/jstsp-ieee>.

Dates:

Manuscript submission: April 1, 2017
First review completed: June 1, 2017
Revised Manuscript Due: July 15, 2017
Second Review Completed: August 15, 2017
Final Manuscript Due: September 15, 2017
Publication: December 2017

Guest Editors:

Nancy F. Chen, Institute for Infocomm Research (I2R), A*STAR, Singapore
Mary Harper, Army Research Laboratory, USA
Brian Kingsbury, IBM Watson, IBM T.J. Watson Research Center, USA
Kate Knill, Cambridge University, UK
Bhuvana Ramabhadran, IBM Watson, IBM T.J. Watson Research Center, USA

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY


www.ieee.org/sp/index.html

FEBRUARY 2017

VOLUME 24

NUMBER 2

ISPLEM

(ISSN 1070-9908)

 A Brief Message From the New Editor-In-Chief <http://dx.doi.org/10.1109/LSP.2016.2635758> *J. E. Fowler* 135

LETTERS

Privacy Leakage of Physical Activity Levels in Wireless Embedded Wearable Systems http://dx.doi.org/10.1109/LSP.2016.2642300	136
..... <i>X. Fafoutis, L. Marchegiani, G. Z. Papadopoulos, R. Piechocki, T. Tryfonas, and G. Oikonomou</i>	
Face Antispoofing Using Speeded-Up Robust Features and Fisher Vector Encoding http://dx.doi.org/10.1109/LSP.2016.2630740	141
..... <i>Z. Boulkenafet, J. Komulainen, and A. Hadid</i>	
Visual Tracking Via Sparse Representation With Reliable Structure Constraint http://dx.doi.org/10.1109/LSP.2016.2645819	146
..... <i>J. Guo, T. Xu, Z. Shen, and G. Shi</i>	
Fusion-Based Variational Image Dehazing http://dx.doi.org/10.1109/LSP.2016.2643168	151
..... <i>A. Galdran, J. Vazquez-Corral, D. Pardo, and M. Bertalmío</i>	
Informed Group-Sparse Representation for Singing Voice Separation http://dx.doi.org/10.1109/LSP.2017.2647810	156
..... <i>T.-S. T. Chan and Y.-H. Yang</i>	
Second-Order Statistics Driven LMS Blind Fractionally Spaced Channel Equalization http://dx.doi.org/10.1109/LSP.2016.2635034	161
..... <i>G. Scarano, A. Petroni, M. Biagi, and R. Cusani</i>	
Successive Hypothesis Testing Based Sparse Signal Recovery and Its Application to MUD in Random Access http://dx.doi.org/10.1109/LSP.2017.2648798	166
..... <i>J. Choi</i>	
Three-Dimensional Integral Imaging for Gesture Recognition Under Occlusions http://dx.doi.org/10.1109/LSP.2016.2643691	171
..... <i>V. J. Traver, P. Latorre-Carmona, E. Salvador-Balaguer, F. Pla, and B. Javidi</i>	
Accurate Cascade Graphic Equalizer http://dx.doi.org/10.1109/LSP.2016.2645280	176
..... <i>V. Välimäki and J. Liski</i>	
Indistinguishability of Compressed Encryption With Circulant Matrices for Wireless Security http://dx.doi.org/10.1109/LSP.2017.2647953 ..	181
..... <i>N. Y. Yu</i>	
On Strategic Multi-Antenna Jamming in Centralized Detection Networks http://dx.doi.org/10.1109/LSP.2017.2649378	186
..... <i>V. S. S. Nadendla, V. Sharma, and P. K. Varshney</i>	
Bridge the Gap Between ADMM and Stackelberg Game: Incentive Mechanism Design for Big Data Networks http://dx.doi.org/10.1109/LSP.2017.2649545	191
..... <i>Z. Zheng, L. Song, and Z. Han</i>	
Joint Power and Admission Control Based on Channel Distribution Information: A Novel Two-Timescale Approach http://dx.doi.org/10.1109/LSP.2017.2651047	196
..... <i>Q. Chen, D. Kang, Y. He, T.-H. Chang, and Y.-F. Liu</i>	

Stability Proof of Iterative Interference Cancellation for OFDM Signals With Blanking Nonlinearity in Impulsive Noise Channels http://dx.doi.org/10.1109/LSP.2017.2650867	<i>M. Mostafa</i>	201
A Novel Noise Jamming Detection Algorithm for Radar Applications http://dx.doi.org/10.1109/LSP.2016.2645793	<i>D. Orlando</i>	206
Construction of Binary Sensing Matrices Using Extremal Set Theory http://dx.doi.org/10.1109/LSP.2016.2638426	<i>R. R. Naidu and C. R. Murthy</i>	211
Total Variation Denoising Via the Moreau Envelope http://dx.doi.org/10.1109/LSP.2017.2647948	<i>I. Selesnick</i>	216
Interpolation of Head-Related Transfer Functions Using Manifold Learning http://dx.doi.org/10.1109/LSP.2017.2648794	<i>F. Grijalva, L. C. Martini, D. Florencio, and S. Goldenstein</i>	221
Robust Time-Delay Estimation for Acoustic Indoor Localization in Reverberant Environments http://dx.doi.org/10.1109/LSP.2016.2642178	<i>J. Choi, J. Kim, and N. S. Kim</i>	226
Recovering Latent Signals From a Mixture of Measurements Using a Gaussian Process Prior http://dx.doi.org/10.1109/LSP.2016.2637312 ..	<i>F. Tobar, G. Rios, T. Valdivia, and P. Guerrero</i>	231
MIMO Beamforming for Secure and Energy-Efficient Wireless Communication http://dx.doi.org/10.1109/LSP.2017.2647982	<i>N. T. Nghia, H. D. Tuan, T. Q. Duong, and H. V. Poor</i>	236
Precoding for Robust Decentralized Estimation in Coherent-MAC-Based Wireless Sensor Networks http://dx.doi.org/10.1109/LSP.2017.2647984	<i>N. K. D. Venkatesowda, B. B. Narayana, and A. K. Jagannatham</i>	240

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



DECEMBER 2016

VOLUME 2

NUMBER 4

ITCIAJ

(ISSN 2333-9403)

PAPERS

Computational Imaging Methods and Models

A Hybrid Markov Random Field/Marked Point Process Model for Analysis of Materials Images http://dx.doi.org/10.1109/TCL.2016.2579601	<i>H. Zhao and M. Comer</i>	395
Plug-and-Play Priors for Bright Field Electron Tomography and Sparse Interpolation http://dx.doi.org/10.1109/TCL.2016.2599778	<i>S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman</i>	408
Mean Squared Error Based Excitation Pattern Design for Parallel Transmit and Receive SENSE MRI Image Reconstruction http://dx.doi.org/10.1109/TCL.2016.2610141	<i>I. Y. Chun, S. Noh, D. J. Love, T. M. Talavage, S. Beckley, and S. J. Kisner</i>	424
Spectral Super-Resolution in Colored Coded Aperture Spectral Imaging http://dx.doi.org/10.1109/TCL.2016.2612943	<i>A. Parada-Mayorga and G. R. Arce</i>	440
Robust Bayesian Target Detection Algorithm for Depth Imaging From Sparse Single-Photon Data http://dx.doi.org/10.1109/TCL.2016.2618323	<i>Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin</i>	456
<i>Computational Image Formation</i>		
Combining Local Regularity Estimation and Total Variation Optimization for Scale-Free Texture Segmentation http://dx.doi.org/10.1109/TCL.2016.2594139	<i>N. Pustelnik, H. Wendt, P. Abry, and N. Dobigeon</i>	468



Computational Imaging Systems

A General Framework for Compressed Sensing and Parallel MRI Using Annihilating Filter Based Low-Rank Hankel Matrix http://dx.doi.org/10.1109/TCL.2016.2601296	<i>K. H. Jin, D. Lee, and J. C. Ye</i>	480
Fast Multilayer Laplacian Enhancement http://dx.doi.org/10.1109/TCL.2016.2607142	<i>H. Talebi and P. Milanfar</i>	496
Spectral CT Reconstruction With Image Sparsity and Spectral Mean http://dx.doi.org/10.1109/TCL.2016.2609414	<i>Y. Zhang, Y. Xi, Q. Yang, W. Cong, J. Zhou, and G. Wang</i>	510
Two-dimensional Autofocus for Spotlight SAR Polar Format Imagery http://dx.doi.org/10.1109/TCL.2016.2612945	<i>X. Mao and D. Zhu</i>	524
Numerical Inversion of Circular arc Radon Transform http://dx.doi.org/10.1109/TCL.2016.2615806	<i>T. A. Syed, V. P. Krishnan, and J. Sivaswamy</i>	540

Computational Imaging Hardware and Software

Vehicle Tracking in Video Using Fractional Feedback Kalman Filter http://dx.doi.org/10.1109/TCL.2016.2600480	<i>H. Kaur and J. S. Sahambi</i>	550
---	----------------------------------	-----

List of Reviewers http://dx.doi.org/10.1109/TCL.2016.2618098		562
---	--	-----

EDICS—Editor's Classification Information Scheme http://dx.doi.org/10.1109/TCL.2016.2622275		564
Information for Authors http://dx.doi.org/10.1109/TCL.2016.2622276		565

2016 INDEX Available online at <http://ieeexplore.ieee.org>

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



DECEMBER 2016

VOLUME 2

NUMBER 4

ITSIBW

(ISSN 2373-776X)

EDITORIAL

Guest Editorial Inference and Learning over Networks <http://dx.doi.org/10.1109/TSIPN.2016.2615526>
..... *V. Matta, C. Richard, V. Saligrama, and A. H. Sayed* 423

PAPERS

Distributed Constrained Recursive Nonlinear Least-Squares Estimation: Algorithms and Asymptotics
<http://dx.doi.org/10.1109/TSIPN.2016.2618318> *A. K. Sahu, S. Kar, J. M. F. Moura, and H. V. Poor* 426

Distributed Detection Over Adaptive Networks: Refined Asymptotics and the Role of Connectivity
<http://dx.doi.org/10.1109/TSIPN.2016.2613682> *V. Matta, P. Braca, S. Marano, and A. H. Sayed* 442

Uniform -Stability of Distributed Nonlinear Filtering Over DNAs: Gaussian-Finite HMMs <http://dx.doi.org/10.1109/TSIPN.2016.2614902> ..
..... *D. S. Kalogerias and A. P. Petropulu* 461

A Distributed Quaternion Kalman Filter With Applications to Smart Grid and Target Tracking
<http://dx.doi.org/10.1109/TSIPN.2016.2618321> *S. P. Talebi, S. Kanna, and D. P. Mandic* 477

Efficient Distributed Online Prediction and Stochastic Optimization With Approximate Distributed Averaging
<http://dx.doi.org/10.1109/TSIPN.2016.2620440> *K. I. Tsianos and M. G. Rabbat* 489

A Decentralized Second-Order Method with Exact Linear Convergence Rate for Consensus Optimization
<http://dx.doi.org/10.1109/TSIPN.2016.2613678> *A. Mokhtari, W. Shi, Q. Ling, and A. Ribeiro* 507

Data Injection Attacks in Randomized Gossiping <http://dx.doi.org/10.1109/TSIPN.2016.2614898>
..... *R. Gentz, S. X. Wu, H.-T. Wai, A. Scaglione, and A. Leshem* 523

Signal Recovery on Graphs: Fundamental Limits of Sampling Strategies <http://dx.doi.org/10.1109/TSIPN.2016.2614903>
..... *S. Chen, R. Varma, A. Singh, and J. Kovačević* 539

Adaptive Least Mean Squares Estimation of Graph Signals <http://dx.doi.org/10.1109/TSIPN.2016.2613687>
..... *P. Di Lorenzo, S. Barbarossa, P. Banelli, and S. Sardellitti* 555

Information Diffusion of Topic Propagation in Social Media <http://dx.doi.org/10.1109/TSIPN.2016.2618324>
..... *S. Mahdizadehghadam, H. Wang, H. Krim, and L. Dai* 569

Metadata-Conscious Anonymous Messaging <http://dx.doi.org/10.1109/TSIPN.2016.2605761>
..... *G. Fanti, P. Kairouz, S. Oh, K. Ramchandran, and P. Viswanath* 582

Evolutionary Information Diffusion Over Heterogeneous Social Networks <http://dx.doi.org/10.1109/TSIPN.2016.2613680>
..... *X. Cao, Y. Chen, C. Jiang, and K. J. Ray Liu* 595

Dual Graph Regularized Dictionary Learning <http://dx.doi.org/10.1109/TSIPN.2016.2605763> *Y. Yankelevsky and M. Elad* 611

On the Difficulty of Selecting Ising Models With Approximate Recovery <http://dx.doi.org/10.1109/TSIPN.2016.2596439>
..... *J. Scarlett and V. Cevher* 625



LIST OF REVIEWERS http://dx.doi.org/10.1109/TSIPN.2016.2618501	639
<hr/>	
EDICS—Editor's Information Classification Scheme http://dx.doi.org/10.1109/TSIPN.2016.2622166	643
Information for Authors http://dx.doi.org/10.1109/TSIPN.2016.2622168	644
<hr/>	
2016 INDEX	<i>Available online at http://ieeexplore.ieee.org</i>



IEEE
Signal Processing Society



General Chairs:

Ph. Coussy, U. Bretagne Sud, FR
W. Gross, McGill U., QC, CA

Program Chairs:

J. Takala, Tampere U. Tech., FI
A. Burg, EPFL, CH

Finance Chairs:

G. Gogniat, U. Bretagne Sud, FR
M. Mattavelli, EPFL, CH

Local Arrangements Chairs:

C. Chavet, U. Bretagne Sud, FR
J. Laurent, U. Bretagne Sud, FR

Publication and publicity chairs:

C. Jégo, Bordeaux INP, FR
V. Lapotre, U. Bretagne Sud, FR

Special sessions chair:

E. Boutillon, U. Bretagne Sud, FR

Industrial liaison chairs:

K. Benkrid, ARM, UK
X. Zhang, SanDisk, US

Demo night chair:

M. Raullet, ATEME, FR

Web site:

K. Martin, U. Bretagne Sud, FR

Asian liaison:

A.-Y. (Andy) Wu, NTU, TW

North-America liaison:

J. Cavallaro, Rice U., TX, US

Sponsored by:

IEEE Signal Processing Society
IEEE Circuits and Systems
Society

SiPS 2017

Preliminary Call for Papers

IEEE Workshop on Signal Processing Systems

Lorient, France, Oct. 3-5 2017

www.sips2017.org



Paper submission deadline (tentative) : May 1st 2017

Acceptance notification (tentative): July 30th 2017

IEEE SiPS is a premier international forum collecting researchers and practitioners from industry and academia for exchanging the latest scientific and technical advances in the area of design and implementation of signal processing systems. It addresses current and future challenges and new directions in research and development of these systems. Prospective authors are invited to submit manuscripts on topics including, but not limited to:

Software Based Design and Implementation of Signal Processing Systems

- Software solutions on programmable digital signal processors and systems
- Application specific instruction-set processor (ASIP) architectures and systems
- SIMD, VLIW and multi-core CPU architectures
- Graphics processing unit (GPU) based massively parallel implementations
- Embedded FPGA architectures

Hardware Based Design and Implementation of Signal Processing Systems

- Low-power signal processing circuits and applications
- High performance VLSI systems
- VLSI designs for 100 Gbps and beyond networking systems
- FPGA and reconfigurable architecture based systems
- System-on-chip and network-on-chip
- VLSI systems for wireless sensor network and RF identification systems

Emerging Technologies

- Vehicular ad hoc networks (VANET)
- Cognitive radio networks
- Internet of Things (IoT)
- Deep learning and reconfigurable/ASIC processors
- Bio-inspired networks
- Context-aware mobile networking
- Wireless body area networks (WBANs)
- Implantable communications
- Tele-medicine/e-health networks

Signal Processing Application Systems

- Audio, speech and language processing
- Biomedical signal processing and bioinformatics
- Image, video and multimedia signal processing
- Information forensics, security and cryptography
- Machine learning for signal processing
- Sensing and sensor signal processing
- Autonomous energy harvesting-based sensor networks
- Signal processing for non-volatile memory systems
- Latency and power constrained signal processing techniques for high-speed networking
- Wireless communications and networking
- Coding and compression
- Multiple-input-multiple-output (MIMO) and communication systems
- Software defined radio

Signal Processing Compensation Techniques for Mixed-Signal Technologies

- Digital compensation techniques for variations in silicon process, temperature, aging
- Error detection and correction for volatile and non-volatile memories
- Power reduction and SNR improvement for on-chip, off-chip interconnects and buses
- Digital compensation signal processing for ADCs, power-amps, MEMS, power controllers

Design Methods of Signal Processing Algorithms and Architectures

- Optimization of signal processing algorithms
- Compilers and tools for signal processing systems
- Algorithm transformation and algorithm-to-architecture mapping
- Dataflow based design methodologies
- Error-tolerant techniques for signal processing

Paper Submission: Authors are invited to submit full-length (max. 6 pages), original, unpublished papers. Previously published papers or papers currently under review for other conferences/journals should not be submitted and will not be considered. Paper format information is available at www.sips2017.org.

IEEE SiPS 2017 has a special theme: “**Reliable signal processing systems**”. Future Signal Processing applications and systems will have to consider reliability in addition to pursuing optimal performance. Submissions that fall into this inter-disciplinary area are particularly encouraged.



CALL FOR PROPOSAL

IEEE International Conference on Multimedia and Expo (ICME) 2019, 2020, and 2021

WWW.IEEE-ICME.ORG

Sponsored By

The IEEE Circuits and Systems Society
The IEEE Communications Society
The IEEE Computer Society
The IEEE Signal Processing Society

This Call for Proposal is distributed on behalf of the Steering Committee for the IEEE International Conference on Multimedia and Expo (ICME) to be held in summer of 2019, 2020 and 2021. IEEE International Conference on Multimedia & Expo (ICME) has been the flagship multimedia conference sponsored by four IEEE societies since 2000. It serves as a forum to promote the exchange of the latest advances in multimedia technologies, systems, and applications from both the research and development perspectives of the circuits and systems, communications, computer, and signal processing communities. An Exposition of multimedia products, animations and industries will be held in conjunction with the conference. ICME 2019, 2020 and 2021 will be the 28th, 29th and 30th meeting in this series.

The conference organizing team is advised to incorporate into their proposal the following items.

- ✧ Proposed Dates (approximately mid-year of 2019 or 2020 or 2021, preferable July)
- ✧ Organizing Committee Members
 - Names for all key positions
 - Biographical information for General Chair(s), Technical Program Chairs, Treasurer/Finance Chair, Workshop Chairs, Industry Chairs, Panel Chairs, Tutorial Chairs, and Special Session Chairs
 - Membership in the **Signal Processing Society**: Multimedia Signal Processing Technical Committee, **Circuits and Systems Society**: Multimedia Systems and Applications Technical Committee, **Communications Society**: Multimedia Communications Technical Committee, and/or **Computer Society**: Technical Committee on Multimedia Computing

(continued on next page)

- ✧ Proposed budget. (For advice on building an IEEE budget please contact sps-conf-proposals@ieee.org)
- ✧ Support that can be anticipated from the local government, universities and or corporations
- ✧ Why select this particular location?
 - Airport and transportation information
 - Customs and Visa regulations
 - Hotel and convention center information (i.e. space diagrams, maps, etc.)
 - Tourist destinations (i.e. museums, natural wonders, etc.)
 - Average weather conditions for the time of year
 - Are there any security concerns? If there is, what procedure will be taken by the OC team?

Submission of Proposal

Proposal for ICME 2019, 2020 or 2021 is currently being accepted. Before submitting a proposal please read the **ICME Charter and Guidelines** on the **ICME website** (<http://www.ieee-icme.org>). Also, please notify the Steering Committee of your intent to propose via an email to Steering Committee Chair Dr. Haohong Wang at haohongwang@gmail.com by **March 1, 2017**. The potential bidders are encouraged to work with a member of ICME steering committee to prepare their bid. If you are interested to solicit Steering Committee feedback for your proposal, please submit a draft proposal no later than **March 15, 2017**. Final proposals should be sent no later than **April 15, 2017**. Please send the proposal to Conference Services Staff, IEEE Signal Processing Society (sps-conf-proposals@ieee.org) and Haohong Wang (haohongwang@gmail.com). As there may be limit of email size, please put large file on a shared drive, e.g., Google Drive, OneDrive, Dropbox, etc.

Proposal Presentation

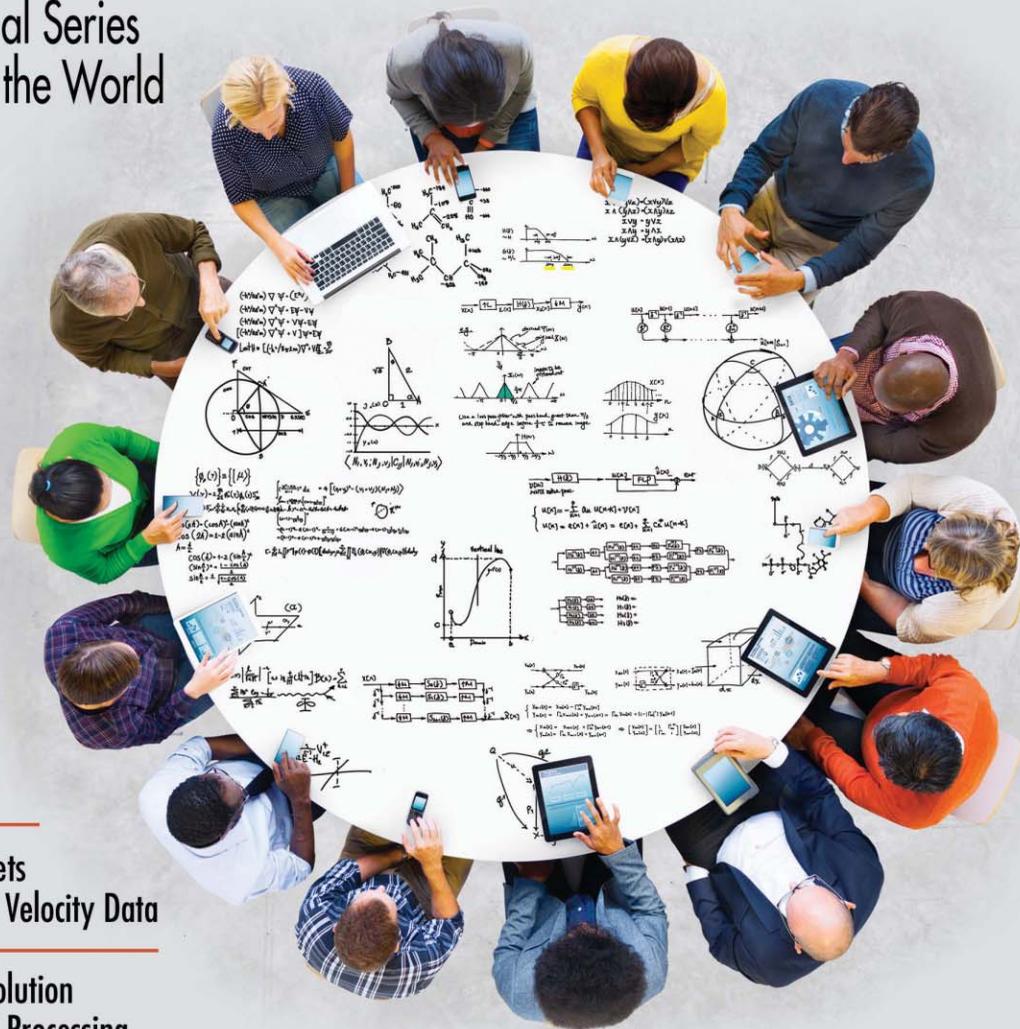
Proposals of interest to the Steering Committee will be asked to present at the Steering Committee meeting at ICME 2017 at Hong Kong. The accepted bid will be required to build an ICME 2019/2020/2021 website at consolidated ICME website within 6 months of proposal acceptance.

IEEE Signal Processing MAGAZINE

Volume 34 | Number 1 | January 2017

HANDS-ON LEARNING AND DESIGN PROJECTS

An Educational Series from Around the World



Compressive Video Sensing

The Technology Behind Personal Digital Assistants

Detecting Exoplanets from Astronomical Velocity Data

A Networking Revolution Powered by Signal Processing



Contents

Volume 34 | Number 1 | January 2017

SPECIAL SECTION

SIGNAL PROCESSING EDUCATION VIA HANDS-ON AND DESIGN PROJECTS

13 FROM THE GUEST EDITORS

Hana Godrich, Arye Nehorai, Ali Tajer, Maria Sabrina Greco, and Changshui Zhang

16 SIGNAL PROCESSING PROJECTS AT TECHNISCHE UNIVERSITÄT DARMSTADT

Tim Schäck, Michael Muma, and Abdelhak M. Zoubir

31 HANDS-ON LEARNING THROUGH RACING

Qing Zhuo, Yanpin Ren, Yongheng Jiang, and Changshui Zhang

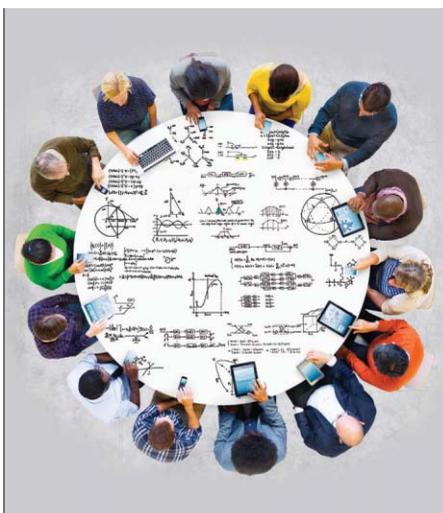
40 TEACHING THE PRINCIPLES OF MASSIVE MIMO

Erik G. Larsson, Danyo Danev, Mikael Olofsson, and Simon Sörman



PG. 116

ADOBE STOCK



ON THE COVER

Integrating more hands-on experiences into formal engineering education is mainstream, and significant efforts are being made in this direction. An insight into the implementation challenges of design projects and experimental platforms from students in their freshmen through senior years, and solutions adopted to address them are offered in this issue of *IEEE Signal Processing Magazine* through a series of article contributions from around the world.

COVER IMAGE: ©ISTOCKPHOTO.COM/RAWPIXEL

FEATURES

52 COMPRESSIVE VIDEO SENSING

Richard G. Baraniuk, Tom Goldstein, Aswin C. Sankaranarayanan, Christoph Studer, Ashok Veeraraghavan, and Michael B. Wakin

67 THE TECHNOLOGY BEHIND PERSONAL DIGITAL ASSISTANTS

Ruhi Sarikaya

COLUMNS

6 Society News

2017 Class of Distinguished Lecturers

9 Special Reports

A Networking Revolution Powered by Signal Processing
John Edwards

48 Perspectives

Small Data, Mid Data, and Big Data Versus Algebra, Analysis, and Topology
Xiang-Gen Xia

82 SP Education

Students' Design Project Series: Sharing Experiences
Hana Godrich

89 Tips & Tricks

An Accurate and Stable Sliding DFT Computed by a Modified CIC Filter
Denis A. Gudovskiy and Lichung Chu

94 Lecture Notes

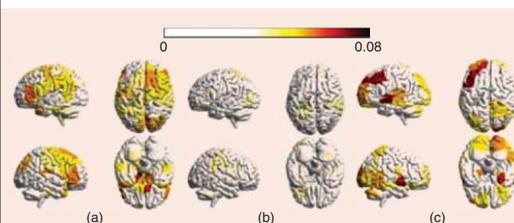
Compressive Privacy: From Information/Estimation Theory to Machine Learning
S.Y. Kung

104 Applications Corner

Discovering New Worlds
Muhammad Salman Khan, James Stewart Jenkins, and Nestor Becerra Yoma

110 Book Digest

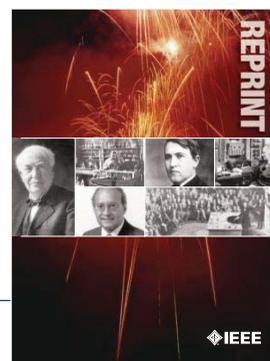
116 In the Spotlight



PG. 82

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\$24.00 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 © 2017 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.2016.2636098



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: _____



2017 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)

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/tsip/tsip-edics/>

The logo for the IEEE Signal Processing Society, featuring the acronym "IEEE" above the full name "Signal Processing Society" in a serif font, with a stylized waveform graphic to the right.

JOIN NOW FOR 2017

The IEEE Signal Processing Society is the world's premier association for signal processing engineers and industry professionals, serving its nearly 17,000 members with highly-ranked publications, world class conferences, professional development resources, job opportunities, and more.

CONNECT

Network with other professionals through SPS conferences, workshops, Technical Committees, Special Interest Groups, and local events curated by more than 180 worldwide SPS Chapters.

SAVE

Access members-only discounts on SPS publications and conferences. Gain eligibility to apply for travel grants to our flagship conferences ICASSP, ICIP, and GlobalSIP.

ADVANCE

Further your career with world-class educational resources, including the new SPS Resource Center, opportunities for awards and recognition, and volunteer opportunities across society activities.



SCAN TO JOIN



@ieeesps



/ieeesps



signalprocessingsociety.org

