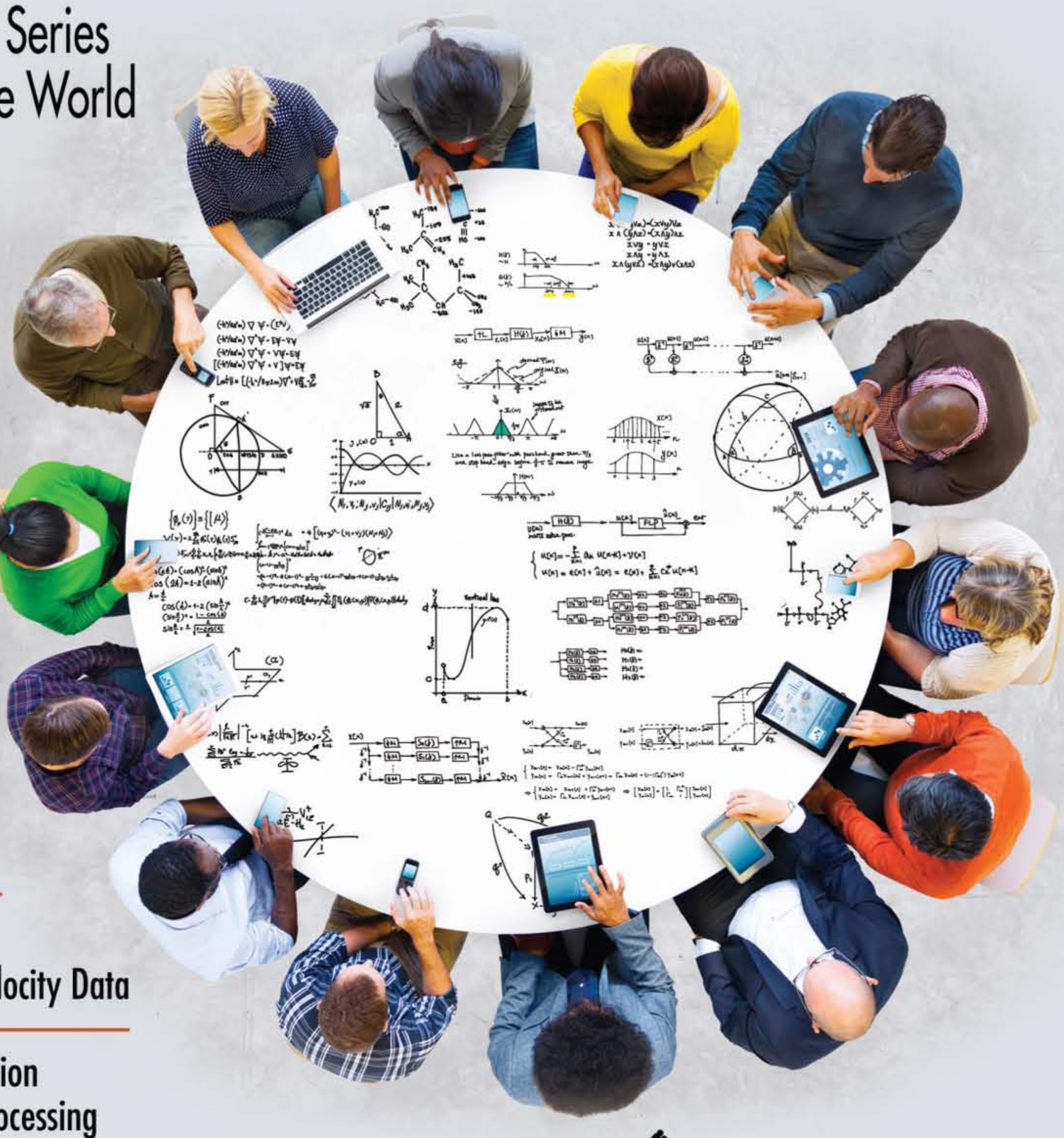
IEEE Signal Processing Magazine

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

IEEE Signal Processing Society

IEEE

Honorary Chairs
Sanjit K. Mitra
Univ. of California at Santa Barbara
Baozong Yuan
Beijing Jiaotong University
General Chair
Xinggang Lin
Tsinghua University
General Co-Chairs
Anthony Vetro
Mitsubishi Electric Research Labs
Min Wu
Univ. of Maryland at College Park
Technical Program Chairs
Jiebo Luo
University of Rochester
Wenjun Zeng
Microsoft Research Asia, China
Yu-Jin Zhang
Tsinghua University
Finance Chair
Yun He
Tsinghua University
Tutorial Chairs
John Apostolopoulos
Cisco
Jiwu Huang
Shenzhen University, China
Special Session Chairs
Pascal Frossard
EPFL
Feng Wu
Univ. of Science and Tech. of China
Awards Chair
Philip Chou
8i, USA
Plenary Chairs
Yao Wang
New York University
Jiangtao Wen
Tsinghua University
Exhibits/Demo Chairs
Qionghai Dai
Tsinghua University
Qi Tian
University of Texas at San Antonio
Industry Chairs
Jian Lu
CCTalk Cloud, China
Debargha Mukherjee
Google
Doctoral Symposium Chair
Dong Xu
University Sydney, Australia
Social Media Chair
Andres Kwansinski
Rochester Institute of Technology
Publicity Chair
Chaker Larabi
Université de Poitiers
Registration Chair
Tiejun Huang
Peking University
Local Arrangement Chairs
Xilin Chen
ICT, Chinese Academy of Sciences
Shengjing Wang
Tsinghua University
Secretary
Zhibo Chen
Univ. of Science and Technology of China
Website Chair
Guijin Wang
Tsinghua University

The 24th IEEE International Conference on Image Processing (ICIP) will be held in the China National Conventional Center, Beijing, China, on 17-20 Septembre 2017. ICIP is the world's largest and most comprehensive technical conference focused on image and video processing and computer vision. The conference will feature world-class speakers, tutorials, exhibits, and a vision technology showcase.

Topics of interest include, but are not limited to:

- Filtering, Transforms, Multi-Resolution Processing
- Restoration, Enhancement, Super-Resolution
- Computer Vision Algorithms and Technologies
- Compression, Transmission, Storage, Retrieval
- Multi-View, Stereoscopic, and 3D Processing
- Multi-Temporal and Spatio-Temporal Processing
- Biometrics, Forensics, and Content Protection
- Biological and Perceptual-based Processing
- Medical Image and Video Analysis
- Document and Synthetic Visual Processing
- Color and Multispectral Processing
- Scanning, Display, and Printing
- Applications to various fields
- Computational Imaging
- Video Processing and Analytics
- Visual Quality Assessment
- Deep learning for Images and Video
- Image and Video Analysis for the Web

Paper Submission:

Authors are invited to submit papers of not more than four pages for technical content including figures and references, with one optional page containing only references. Submission Instructions, templates for the required paper format, and information on "no show" policy are available at 2017.ieeeicip.org.

Tutorials, Special Sessions, and Challenge Sessions Proposals:

Tutorials will be held on September 17, 2017. Tutorial proposals must include title, outline, contact information, biography and selected publications for the presenter(s), and a description of the tutorial and material to be distributed to participants. For detailed submission guidelines, please refer to the tutorial proposals page. Special Sessions and Challenge Session Proposals must include a topical title, rationale, session outline, contact information, and a list of invited papers/participants. For detailed submission guidelines, please refer the ICIP 2017 website at 2017.ieeeicip.org.

Important Dates

Special Session Proposal:	Nov. 15, 2016
Notification of Special Session acceptance:	Dec. 15, 2016
Tutorial Proposal:	Dec. 15, 2016
Notification of Tutorial accptance:	Jan. 15, 2017
Full Paper & Special Session Submission:	Jan. 31, 2017
Notification of acceptance:	Apr. 30, 2017
Camera-Ready Papers:	May. 31, 2017

Digital Object Identifier 10.1109/MSP.2016.2636083

IEEE | IEEE Signal Processing Society | IEEE International Conference on Image Processing

Contents

Volume 34 | Number 1 | January 2017

SPECIAL SECTION

SIGNAL PROCESSING EDUCATION VIA HANDS-ON AND DESIGN PROJECTS

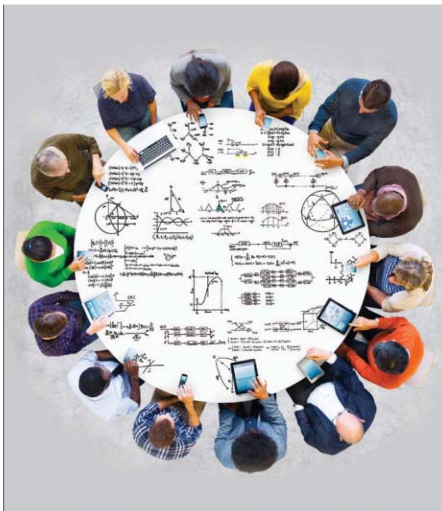
13 FROM THE GUEST EDITORS
Hana Godrich, Arye Nehorai, Ali Tajer, Maria Sabrina Greco, and Changshui Zhang

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

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

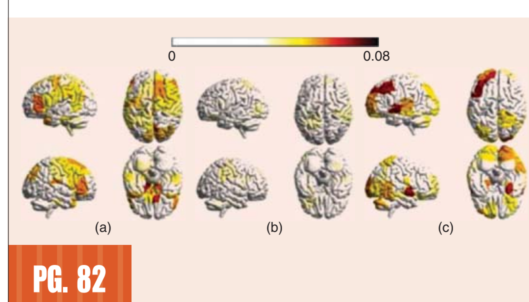
89 Tips & Tricks
An Accurate and Stable Sliding DFT Computed by a Modified CIC Filter
Denis A. Gudovskiy and Lichung Chu

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



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.2016.2636098

DEPARTMENTS

3 **From the Editor**
Signal Processing: The Expected and the Unexpected
Min Wu

4 **President's Message**
Volunteerism Makes a Positive Difference
Rabab Ward

113 **Dates Ahead**

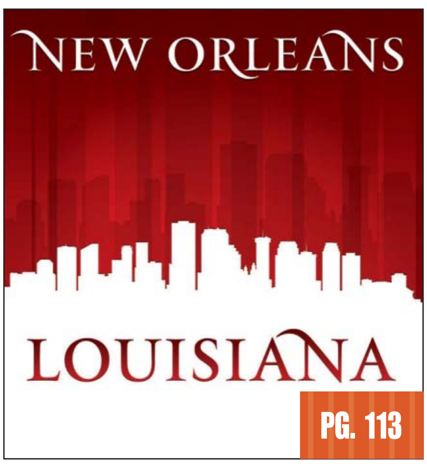


IMAGE LICENSED BY INGRAM PUBLISHING

ICASSP 2017 will be held in New Orleans, Louisiana, 5–9 March.

SUSTAINABLE FORESTRY INITIATIVE Certified Chain of Custody Promoting Sustainable Forestry www.sfiprogram.org SFI-01681

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

Digital Object Identifier 10.1109/MSP.2016.2636099

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

Signal Processing: The Expected and the Unexpected

It will be the start of another new year when you receive this issue of *IEEE Signal Processing Magazine* (*SPM*). Happy 2017 to all our readers, editors, and reviewers!

Not long before I began writing this editorial, the 2016 edition of the IEEE International Conference on Image Processing (ICIP) was successfully held in Phoenix, Arizona. In addition to the rich and timely technical sessions that ICIP is well known for, the ICIP 2016 team—led by General Chair Prof. Lina Karam, who is also serving on *SPM*'s senior editorial board, and Industrial Program Chair Dr. Haohong Wang—spearheaded the first Visual Innovation Award. Going over the finalists' roster, you may very well find yourself having been a user of some of these technologies: the YouTube video streaming service, the Lytro light-field camera, the Intel RealSense camera technology, the CUDA high-performance computing by NVIDIA, the Netflix movie streaming service, the Oculus virtual reality technology, and the Microsoft Kinect.

For many signal processing professionals, including those who regularly attend ICIP—a flagship conference of the IEEE Signal Processing Society (SPS)—it might almost have been taken for granted that signal processing plays a key role behind these visual innovations. Whether it is image formation, sensing, compression, or communications, signal processing provides the underlying technical foundation.

Digital Object Identifier 10.1109/MSP.2016.2632218
Date of publication: 11 January 2017

Right after ICIP, I briefly stopped in the San Francisco Bay area, where I gave a keynote speech at a North American alumni forum of my college alma mater, Tsinghua University, in Beijing, China. Different from ICIP, I did not expect this forum to be a venue to see so much signal processing other than the talk on microsignals for media security that I would be giving. I did my undergraduate study in the Department of Automation at Tsinghua University. "Automation" as an engineering major covers a combination of control and robotics, electronic sensing and diagnosis, signal processing, and pattern recognition; within the department, different specialty directions were rather compartmentalized historically. Perhaps it was due to the difficulty to find an exact matching department in North American universities that college alumni from the department went in different ways when pursuing their graduate studies in North America. Among them, you will find experts on securing sensors and sensor network, on supply chain management behind some of the most wanted consumer products, on designing the next-generation mass spectrometry analyzer, and on international finance and policy making, just to name a few.

Yet through this stimulating day-long event, I learned a great deal about many broad applications of signal processing. For example, a keynote speech given before mine provided an overview on designing and analyzing sensing signals for fault-tolerant operations in such complex systems as the quality control and enhancement in steel manufacturing and the signaling in China's high-speed train systems. As it turned out, many challenging issues addressed by the keynote speech have benefited from signal processing theories and techniques. Two panel discussions on the recent hype of artificial intelligence and the Internet of Things also touched on such issues as sensing, denoising, and statistical learning from signals and data. In addition, several alumni who are successful in venture capital investment highlighted the important roles of data and data analytics that they saw in developing sustainable new businesses.

Most speakers at the alumni forum would not consider themselves to be professionals in signal processing, and not many have read our magazine. It reminded me of "Signal Processing Inside," a notion coined in *SPM*'s September 2004 editorial by then Editor-in-Chief Prof. K.J. Ray Liu, and the blurred boundaries between disciplines discussed in my September 2016 editorial. Inspired by those expected and unexpected venues where signal processing shines, I am working with our magazine editors to develop leads on informative articles for our readers in the coming months. We welcome your suggestions on topics that you would like to read about.

Best wishes to you all for a prosperous new year ahead—another year filled with exciting signal processing!

SP

PRESIDENT'S MESSAGE

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

Volunteerism Makes a Positive Difference

Happy New Year! The onset of a new year is an exciting time filled with renewal, opportunity, and possibility. As we close the book on 2016, it's only natural to reflect on our progress, celebrate our successes, and learn from our failures. But the beginning of a new year is also empowering—it reignites your fire to try something new, to set goals, to be open to change and embrace new challenges. The IEEE Signal Processing Society (SPS) hopes to invoke these principles and nimbleness into 2017 and for many years to come. We hope that you'll continue to work with us through this exciting time!

I would be remiss to not mention that 2017 marks the expiration of a valued member of the SPS Executive Committee's term. In 2013, the SPS Executive Committee divided the role of vice president, Awards and Membership, establishing a newly individualized role of vice president, Membership, to better examine, understand, and serve SPS's growing and diverse membership. Dr. Kostas Plataniotis was chosen as to fill this position. His vision and innovation were essential to not only proving the role's necessity but to driving its purpose and setting the course for future member services and activities. We thank him for his effective and incredible service and are excited to welcome Dr. Nikos Sidiropoulos to his new position at SPS vice president, Membership.

This past year brought about a lot of exciting changes—in June, we launched the new SPS website, accompanied by the SPS Resource Center; http://rc.signalprocessingsociety.org. With these, a new era began for the Society, looking toward and acting on a rapidly changing future. Our organization is changing. Our Society is changing. And, most notably, our membership is changing, and we have to find new and effective ways to keep members of all ages, career stages, professions, and fields engaged and involved with the SPS and its activities. Volunteerism is a great way to encourage early involvement, building loyalty across a diverse member body with varying interests, availability, needs, and expertise. Volunteerism has evolved in itself, and SPS is working toward building a strong volunteer base, with roles to suit the growing needs for the Society and its members alike.

The nature of "volunteering" used to be immersive and intimidating, with excessive time commitments that made volunteerism seem like too much to juggle among other activities. Now, with the help of technology and the evolution of workplaces, a volunteer has new flexibility and freedom to choose his/her level of involvement, from demanding high-profile leadership roles to "microvolunteering" opportunities that spark interest and action without consuming as much from our already busy schedules. Over the past couple of years, the SPS has expanded its volunteer roles to encourage diverse involvement opportunities. You can choose a position that best fits your lifestyle whether you're a student, young professional, in the middle of your career, or your career is winding down—or even if you are retired but want to stay active.

The SPS relies on its dedicated volunteer base—more than 1,000 members strong—to develop and manage Society activities, products, and services. Decision makers who sit on our boards and committees play an integral role in the Society and its operations. These high-level roles cover a wide array of Society needs in the areas of conferences, publications, membership, education, and more. Sitting on boards and committees, while time-consuming, can be incredibly rewarding and prestigious. Many high-ranking board members move on to become decision makers in broader-scale positions within the IEEE.

Even among publications, conferences, education, and membership, there are a multitude of opportunities of varying levels of involvement. You can be a reviewer or editor of one of our Society publications, form a committee to propose and host a conference or a meeting in a desired area, or propose a seasonal school workshop near you. Want to get involved locally? Form an SPS Chapter, or attend an event of an existing local Chapter to strengthen connections with other signal processing professionals

Digital Object Identifier 10.1109/MSP.2016.2628418
Date of publication: 11 January 2017

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

near you. Why not host a networking event or technical talk in conjunction with a local Chapter and involve local industry partners?

The SPS has several technical committees (TCs) and special interest groups (SIGs) that help steer the technical direction of the Society, contributing their expertise in regards to SPS conferences, awards, publications, and educational activities. Anyone can become an affiliate member of a TC or SIG—both are a great way to not only get involved with important Society activities but to build relationships with other SPS members who share similar technical interests to you.

Many of us don't have time to dedicate to planning events, serving on boards, or reviewing papers. Maybe you're new to the Society and want to get more involved but don't really know where you can step in. The SPS is always looking for volunteers to help with our ongoing visibility efforts. This can entail something as simple as sending out a quick Tweet to promote the SPS or share signal processing news, or it can be as involved as writing a post for our new SPS blog; http://signalprocessingsociety.org/publications-resources/blog. We even have a group of volunteers on call when signal processing sources are needed for external media stories. Social media ambassadorship and blog contributions are both great ways for younger members to get involved early without yet committing to more serious roles.

Volunteering for the SPS in any capacity—whether for a couple of hours a month to several days a year—is a great way to get involved, build your resume, make connections within the field, and expand your career options. These are just a sampling of the many ways SPS members can have a hand in Society activities and another way the SPS strives to help its members reach their goals—whether the goals are only the new year or beyond.

The SPS wishes you a happy, healthy, and prosperous new year. For full information about volunteering with the SPS, visit our website at http://signalprocessingsociety.org. If you have questions or need guidance, please feel free to contact me or our SPS Membership and Content Administrator Jessica Perry at jessica.perry@ieee.org.

SP

Digital Object Identifier 10.1109/MSP.2016.2581299

2017 Class of Distinguished Lecturers

The IEEE Signal Processing Society's (SPS's) Distinguished Lecturer Program provides the means for Chapters to have access to well-known educators and authors in the fields of signal processing to lecture at Chapter meetings. While many IEEE Societies have similar programs, the SPS provides financial support for the Chapters to take advantage of this service. Chapters interested in arranging lectures by the Distinguished Lecturers can obtain information from the Society's web page (http://signalprocessingsociety.org/professional-development/distinguished-lecturer-program) or by sending an e-mail to sp.info@ieee.org.

Candidates for the Distinguished Lecturer Program are solicited from the Society technical committees, editorial boards, Chapters, and other boards and committees by the Awards Board. The Awards Board vets the nominations, and the Board of Governors approves the final selection. Distinguished Lecturers are appointed for a term of two calendar years. Distinguished Lecturers named for 2017 are as follows.

Vivek K. Goyal

Vivek K. Goyal obtained his B.S. degree in mathematics (1993) and his B.S.E. degree in electrical engineering (1993) from the University of Iowa, where he received the John Briggs Memorial Award for the top undergraduate across all colleges. He obtained the M.S. degree (1995) and Ph.D. degree (1998) in electrical engineering from the University of California, Berkeley, where he received the Eliahu Jury Award for outstanding achievement in systems, communications, control, or signal processing.

Chapters interested in arranging lectures by the Distinguished Lecturers can obtain information by sending an e-mail to sp.info@ieee.org.

Dr. Goyal was a member of technical staff in the Mathematics of Communications Research Department, Bell Laboratories, Lucent Technologies, (1998–2001) and a senior research engineer for Digital Fountain, Inc. (2001–2003). He was the Esther and Harold E. Edgerton Associate Professor of Electrical Engineering, Massachusetts Institute of Technology (2004–2013), adviser, 3dim Tech, Inc. (winner of the 2013 MIT $100K Entrepreneurship Competition Launch Contest Grand Prize and 2013 MassChallenge Accelerator Gold), and was subsequently with Nest, an Alphabet company (2014–2016). He is now with the Department of Electrical and Computer Engineering of Boston University.

He is an IEEE Fellow and was awarded the IEEE SPS Magazine Award (2002), the IEEE SPS Best Paper Award (2014), and a National Science Foundation (NSF) CAREER Award. The work he supervised won student best paper awards at the IEEE Data Compression Conference in 2006 and 2011 and the IEEE Sensor Array and Multichannel Signal Processing Workshop in 2012 as well as five MIT thesis awards. He is a coauthor of *Foundations of Signal Processing* (Cambridge University Press, 2014).

Dr. Goyal served on the IEEE Image and Multidimensional Signal Processing Technical Committee (2003–2009); IEEE Image, Video, and Multidimensional Signal Processing Technical Committee (2014); and the steering committee of *IEEE Transactions on Multimedia* (2013). He has served as editorial board member, *Foundations and Trends and Signal Processing* (2006–present); scientific advisory board of the Banff International Research Station for Mathematical Innovation and Discovery (2011–present); the IEEE SPS Computational Imaging SIG (2015–present); the IEEE Standing Committee on Industry DSP Technology (2016–present); technical program cochair, International Conference on Sampling Theory and Application (2015); and conference cochair, SPIE Wavelets and Sparsity conference series (2006–2016).

Digital Object Identifier 10.1109/MSP.2016.2622958
Date of publication: 11 January 2017

Dr. Goyal's research interests include computational imaging, human perception, decision making, sampling, quantization, and source coding theory. His lecture topics include first-photon imaging and other extreme optical imaging, social learning in decision-making groups, teaching signal processing with geometry, and the optimistic Bayesian: replica method analysis of compressed sensing.

Christine Guillemot

Christine Guillemot holds a Ph.D. degree from Ecole Nationale Superieure des Telecommunications Paris. She was with FRANCE TELECOM, where she was involved in various projects in the area of coding for TV, high-definition TV, and multimedia (November 1985 to October 1997) and she worked at Bellcore, New Jersey, as a visiting scientist (January 1990 to mid-1991). Since November 1997, she has been the director of research at INRIA, as the head of a research team dedicated to the design of algorithms for the image and video processing chain, with a focus on analysis, representation, compression, and editing, including for emerging modalities such as high dynamic range imaging and light fields.

Dr. Guillemot has coauthored nine book chapters, 65 publications in peer-reviewed international journals (*IEEE Transactions on Signal Processing, IEEE Transactions on Image Processing, IEEE Transactions on Information Theory,* and *IEEE Transactions on Circuits and Systems for Video Technology*), 162 publications in international conferences [IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), IEEE International Conference on Image Processing, IEEE International Workshop on Multimedia Signal Processing (MMSP), and European Signal Processing Conference (EUSIPCO)], and 24 granted patents.

Dr. Guillemot is an IEEE Fellow. She has served as associate editor, *IEEE Transactions on Image Processing* (2000–2003 and 2014–2016); associate editor, *IEEE Transactions on Circuits and Systems for Video Technology* (2004–2006); associate editor, *IEEE Transactions on Signal Processing* (2007–2009); associate editor, *IEEE Journal on Selected Topics in Signal Processing* (2013–2015); member, IEEE Image and Multidimensional Signal Processing Technical Committee (2001-2006); member, IEEE Multimedia Signal Processing Technical Committee (2005–2008); member, IEEE Image, Video, and Multidimensional Signal Processing Technical Committee (2013–present); senior area editor, *IEEE Transactions on Image Processing* (2016–2017); and steering committee member, *IEEE Transactions on Multimedia* (2016).

Over the past 20 years, Dr. Guillemot's research has focused on numerous aspects of image and video processing: modeling, representation, compression, and communication. Her contributions concern algorithms for image and video analysis, representation, coding, communication, and for inverse problems such as superresolution, inpainting, and restoration. Her lecture topics include sparsity and dimensionality reduction in image compression and superresolution; multiview and light fields processing: from analysis, representation, compression to rendering; and from image to video and multiview inpainting.

Petros Maragos

Petros Maragos received the M.Eng. diploma in electrical engineering from the National Technical University of Athens (NTUA) in 1980 and the M.Sc. and Ph.D. degrees from Georgia Tech, Atlanta, in 1982 and 1985. In 1985, he joined the faculty of the Division of Applied Sciences at Harvard University, Cambridge, Massachusetts, where he worked for eight years as professor of electrical engineering, affiliated with the Harvard Robotics Lab. In 1993, he joined the faculty of the School of Electrical and Computer Engineering (ECE) at Georgia Tech, affiliated with its Center for Signal and Image Processing. From 1996 to 1998, he had a joint appointment as director of research at the Institute of Language and Speech Processing in Athens. Since 1999, he has been working as a professor at the NTUA School of ECE, where he is currently the director of the Intelligent Robotics and Automation Lab. He has held visiting scientist positions at the Massachusetts Institute of Technology in the fall of 2012 and at the University of Pennsylvania in the fall of 2016.

Prof. Maragos served as associate editor, *IEEE Transactions on Acoustics, Speech, and Signal Processing* (1989–1990); and *IEEE Transactions on Pattern Analysis and Machine Intelligence*; general chair, IEEE International Conference on Visual Communications and Image Processing (1992); general chair, International Symposium on Mathematical Morphology and Its Applications to Image/Signal Processing (1996); general chair, MMSP (2007); program chair, European Conference on Computer Vision (2010); ECCV Workshop on Sign, Gesture, and Activity (2010); Dagstuhl Symposia on Shape (2011 and 2014); Intelligent Robots and Systems Workshop on Cognitive Mobility Assistance Robots (2015); general chair, EUSIPCO (2017); member, SPS Digital Signal Processing Technical Committee (1992–1998); IEEE SPS Image and Multidimensional Signal Processing Technical Committee (1995–1999); IEEE SPS Multimedia Signal Processing Technical Committee (2009–2012); and member, Greek National Council for Research and Technology.

He is the recipient or corecipient of several awards for his academic work, including: U.S. NSF Presidential Young Investigator Award (1987–1992); IEEE SPS Young Author Best Paper Award (1988), IEEE SPS Best Paper Award (1994), IEEE W.R.G. Baker Prize Award for the most outstanding original paper (1995), Pattern Recognition Society's Honorable Mention Best Paper Award (1996), and Best Paper Award, Conference on Computer Vision and Pattern Recognition-2011 Workshop on Gesture Recognition.

Prof. Maragos was elected IEEE Fellow for his research contributions in 1995 and received the 2007 EURASIP Technical Achievements Award for contributions to nonlinear signal processing, systems theory, image, and speech processing. In 2010, he was elected fellow of EURASIP for his research contributions. He has been elected IEEE SPS Distinguished Lecturer for 2017–2018.

Prof. Maragos' research and teaching interests include signal processing, systems theory, machine learning, image processing and computer vision, audio and speech/language processing, cognitive systems, and robotics. In the aforementioned areas he has published numerous papers, book chapters, and has also coedited three Springer research books, one on multimodal processing and two on shape analysis.

Prof. Maragos' lecture topics include multimodal spatiotemporal signal processing and audio-visual perception, nonlinear signal processing and dynamical systems on lattices, morphological and variational methods in image analysis and computer vision, graph-based methods for clustering and segmentation, and multimodal gesture and spoken command recognition in human-robot interaction.

Athina P. Petropulu

Athina P. Petropulu received her undergraduate degree from the National Technical University of Athens, Greece, in 1986, and the M.Sc. and Ph.D. degrees from Northeastern University, Boston, Massachusetts, in 1988 and 1991, respectively, all in electrical and computer engineering (ECE). Since 2010, she has been a professor in the ECE Department at Rutgers University, New Brunswick, New Jersey, having served as chair of the department during 2010–2016. Before that she was a member of faculty at Drexel University, Philadelphia, Pennsylvania.

Dr. Petropulu is an IEEE Fellow (2008) and the recipient of the 1995 Presidential Faculty Fellow Award given by NSF and the White House. She has served as editor-in-chief, *IEEE Transactions on Signal Processing* (2009–2011); IEEE SPS vice president, conferences (2006–2008); member-at-large, IEEE SPS Board of Governors (2004–2005); general chair, ICASSP 2005; recipient, IEEE Signal Processing Magazine Best Paper Award (2005); recipient, IEEE SPS Meritorious Service Award (2012); member, IEEE SPS Fellow Reference Committee (2012–2014); and was selected as an IEEE Distinguished Lecturer for the SPS (2017–2018).

Dr. Petropulu's research interests span the area of statistical signal processing, wireless communications, signal processing in networking, physical layer security, and radar signal processing. Her research has been funded by various government industry sponsors including the NSF, the Office of Naval research, the U.S. Army, the National Institutes of Health, the Whitaker Foundation, and Lockheed Martin. Her lecture topics include sparse sensing-based multiple-input, multiple-output radars; the coexistence of radar and communication systems; cooperative approaches for physical layer security; cooperative approaches for improving the performance of wireless networks; mobile beamforming; and localization of brain activations based on electroencephalogram recordings and sparse signal recovery theory.

Brian M. Sadler

Brian M. Sadler is the U.S. Army senior research scientist for Intelligent Systems, at the Army Research Laboratory (ARL) in Maryland. He received his undergraduate and master's education in electrical engineering from the University of Maryland in 1984 and the Ph.D. degree in electrical engineering from the University of Virginia in 1993.

He was an associate editor of *IEEE Transactions on Signal Processing* (1999–2001, 2008–2009, 2015–present), EURASIP *Signal Processing*, and *IEEE Signal Processing Letters* (2006–2007). He has been a guest editor of several journals including *IEEE Journal of Selected Topics for Signal Processing*, *IEEE Journal on Selected Areas in Communications,* and *IEEE Signal Processing Magazine*; lead guest editor, *International Journal of Robotics Research* special issue on networked robotics; general cochair, IEEE Global Conference on Signal and Information Processing (GlobalSIP 2016); member, SPS Sensor Array and Multichannel Technical Committee (2006–2011 and 2015–present); member, Signal Processing for Communications Technical Committee (1999–2005); and cochair, IEEE Robotics and Automation Society Technical Committee on Networked Robotics.

Dr. Sadler received the IEEE SPS Best Paper Award in 2006 and 2010, several ARL awards, three Army R&D Achievement awards, as well as the Outstanding Invention of the Year Award from the University of Maryland in 2008.

Dr. Sadler is a Fellow of the IEEE and ARL, and he has lectured at the Johns Hopkins University Whiting School of Engineering for 14 years.

Dr. Sadler's research interests span intelligent systems, with an emphasis on distributed collaborative operation, including multiagent autonomy, cognitive networking, distributed sensing and signal processing, and mixed-signal circuit architectures for low power sensing and cognition. His recent work focuses on collaborative physical agents in stressful and complex environments; "20-questions" strategies for machines to query humans; and the combination of distributed computation, control, and cognitive networking. His lecture topics include distributed collaborative intelligent systems, human-autonomy querying and interaction, and autonomous networking.

SP

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

John Edwards

A Networking Revolution Powered by Signal Processing

Data networks are extending their reach into virtually every corner of life. The rapidly emerging Internet of Things (IoT), for instance, promises to bring connectivity to just about everything. Technology analysis firm Gartner (http://www.gartner.com/newsroom/id/3165317) predicts that by 2020 there will be over 20 billion connected devices.

Signal processing is helping the IoT and other network technologies to operate faster, more efficiently, and very reliably. Advanced research also promises to open new opportunities in key areas, such as highly secure communication and various types of wireless networks.

Seeking quantum communication

At Stanford University, researchers are examining quantum communication as a potential way to quickly and reliably secure Internet traffic. Yet before this goal can be reached, they will have to overcome several major technical challenges, including developing devices that can actually send and receive quantum data. An important step in creating such devices is to develop a quantum light source that might someday serve as the basis for secure quantum data transfers.

Ordinary lasers can't be used for secure communication since they emit a "classical" light that would enable unauthorized parties to extract data without detection. A secure quantum network, on the other hand, would be based on quantum light in which a single unit of light—a lone photon—could not even be measured without being destroyed. Therefore, an efficient quantum light source would allow completely secure communication.

A research team led by Jelena Vuckovic, a Stanford professor of electrical engineering, has spent the past several years working toward the development of nanoscale lasers and quantum technologies that might someday enable conventional computers to communicate faster and more securely using light instead of electricity. Vuckovic and her team, including Kevin Fischer, a doctoral candidate and lead author of a paper describing the project, believe that a modified nanoscale laser can be used to efficiently generate quantum light for fully protected quantum communication. "Quantum networks have the potential for secure end-to-end communication wherein the information channel is secured by the laws of quantum physics," Fischer says.

"Our quantum light source produces single photons, one at a time, on demand," Fischer continues. "Our technology also poses the potential for two or three photon sources as well." Such light sources will be critical for future quantum networking and computation applications. "They serve as the signal that goes into the input of any quantum processor," Fischer notes.

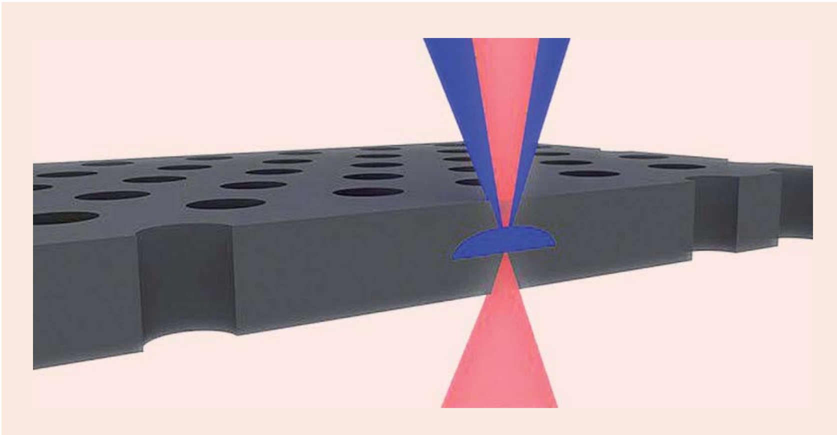
Optical signal processing is handled by the optical elements themselves at the speed of light. Quantum light, Fischer says, is effectively the study of signals that are represented by continuous random variables. "Thus, in order to characterize our quantum signals, we borrow a variety of techniques from classical signal processing," he notes. "Some examples … are Fourier analysis to analyze the spectral content of our signals and frequency filtering to examine specific spectral content."

The biggest challenge the researchers have faced so far is dealing with the fact that quantum light is far weaker than the rest of the light emitted by a modified laser, making it difficult to detect. Addressing this obstacle, the team developed a method to filter out the unwanted light, enabling the quantum signal to be read much better. "Some of the light coming back from the modified laser is like noise, preventing us from seeing the quantum light," Fischer says. "We canceled it out to reveal and emphasize the quantum signal hidden beneath."

To deal with the noise issue, the researchers turned to self-homodyning—an interferometric technique that was originally invented as a method for detecting radio-frequency signals, mixing the signal in question with a strong local oscillator. "We used an optical analog of this technique to isolate quantum as opposed to classical signals. By carefully adjusting how the canceling light and the classical light overlap, the unwanted light is canceled and the once-hidden quantum light is revealed," he says (Figure 1).

Self-homodyning and interferometric techniques generally require precise

Digital Object Identifier 10.1109/MSP.2016.2616376
Date of publication: 11 January 2017



YOUSIF KELAITA/STANFORD UNIVERSITY

FIGURE 1. An enlarged artist's rendering showing a gallium arsenide chip. The pink vector (at the bottom) depicts "classical" or laser light entering the chip. The blue structure in the center is indium arsenide. This material acts like a special filter that allows classical light to pass through while also generating quantum light (shown in blue) that provides a secure way to transmit data.

phase-locking of the signal and oscillator fields, which is challenging to achieve with light compared to radio frequency fields. "This is challenging with light, because the wavelength is much smaller—on the order of microns as opposed to meters—which then necessarily requires greater precision," Fischer says. "Therefore, our advance was to find a device structure that generated both the local oscillator field and the quantum signal, which were then inherently aligned to one another."

The biggest technical challenge still facing the researchers is scaling the optical light devices down to a size that will allow for integration into quantum networks. "We are working toward demonstrating adapting this technique in an on-chip quantum network, where light propagates down a waveguide as opposed to through free-space," Fischer comments.

Using state-of-the-art nanofabrication technology, the team is currently engineering its first quantum light devices.

Building such devices with the required low tolerances challenges even the most advanced fabrication techniques. "Our interferometer's largest critical dimension is microns and smallest critical dimension is nanometers," Fischer says.

Creating a practical and cost-effective approach to integrating optical light devices into standard complementary metal–oxide–semiconductor (CMO) fabrication processes is yet another challenge facing the researchers. "Therefore, we're also investigating CMOS-compatible material platforms that can support our technology," Fischer says.

As the researchers turn their attention toward developing a functional prototype, commercial applications exist as only a distant possibility. "Not yet," Fischer says. "We first need to demonstrate that our device works in a waveguide-based system."

Simpler sensor networks

Ioannis Schizas, an assistant professor of electrical engineering at the University of Texas at Arlington, is developing a sensing environment that would use multiple simple devices to collect and process data that currently requires the power of a supercomputer (Figure 2). "Sensors provide huge amounts of data, but using and applying the data they collect requires a very powerful computer," Schizas says. "I hope to eliminate that need through simplicity of design."

As he creates the new sensing environment, Schizas is using several different types of commonly available sensors to collaborate with each other and gather various types of data that can be either sorted or ignored. He hopes to eliminate the need for supercomputing by using optimization techniques to determine the best placement of sensors, including thermometers, accelerometers, pressure sensors, and acoustic sensors equipped with digital signal processors (DSPs) and wireless communications support. Schizas says his research relies on the development of novel signal processing techniques. "It is fair to say this is a signal processing research project," he states.

Schizas says his research is currently focused on the development of general

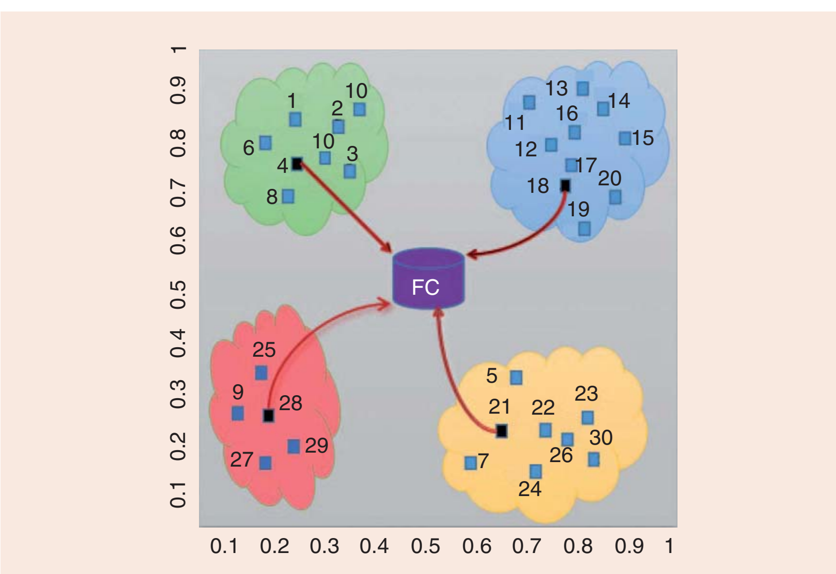


FIGURE 2. A network of sensors observing a field. Ioannis Schizas, an assistant professor of electrical engineering at the University of Texas at Arlington, is developing a sensing environment that would use multiple simple devices to collect and process data that currently requires the power of a supercomputer. (Photo courtesy of the University of Texas at Arlington.)

algorithms with learning capabilities that can identify different informative portions within various types of sensor data that may adhere to different data models. "Heterogeneous sensing systems consist of sensors with different types of sensing and communication capabilities," he explains. "The main challenge is that the often large amount of acquired raw sensed data doesn't provide any clue of what lies beneath the sensed field."

One of the project's major goals is to cluster data into groups containing specific information about different sources or entities of interest. "Distributed processing techniques that do not require a central processing center are also being developed," Schizas says.

Schizas is using canonical correlation analysis (CCA) to reveal correlated data that contains similar information content. "Further, norm-one regularization is combined with CCA to identify the specific entries that have similar information content and perform clustering," he says. The proposed framework is solved using a block coordinate descent approach; principal component analysis is employed to determine the number of underlying sources/objects of interest. "Further, moving averaging and least-mean squares filtering are employed to perform denoising and signal reconstruction," he notes.

The project promises to open new ways for data-driven data clustering where there is no need to rely on available statistical data models. Learning algorithms are being developed that solely rely on the available data to perform information clustering. "The proposed framework is pretty flexible, and it is expected to create benefits in many areas, including target tracking, machine learning, and image processing," Schizas says.

Schizas hopes that the project will eventually lead to self-organizing sensor networks incorporating a variety of positive attributes, including low energy demands, robust architectures, and prolonged life expectancies, deployed in fields such as health care, defense, and structural and other types of monitoring. "Especially in applications involving environmental monitoring and the processing of ecological and climatic data, the project will introduce beneficial data mining solutions to deal with the heterogeneous and high volume data," he notes.

"Most, if not all, of the challenges encountered so far relate to signal processing issues," Schizas says. A current concern is finding a way to deal with sensor data that contains information for multiple objects of interest. "This gives rise to overlapping information clusters that are well known to challenge all existing clustering techniques," Schizas says.

Schizas is satisfied with the progress made to date. "So far, we have developed a novel combination of CCA with principal component analysis to identify sensor data that contains information about multiple sources and determine in that way the overlapping information clusters," he says. Yet the current approach works only for linear models, not for nonlinear data models. "Our goal is to generalize our framework to address the nonlinear case," Schizas says. "Further, the presence of nonstationary and time-varying statistics is another challenge that we are currently trying to address, relying on online and adaptive learning."

Schizas notes that the project is still relatively new and that much work still remains to be done. "There is no commercial interest yet, but as we improve upon computational complexity and generalize the applicability of the proposed algorithms we expect to permeate benefits in existing sensing systems and raise commercial interest," he says.

Collision-free Wi-Fi

Researchers in the Massachusetts Instute of Technology's (MIT's) Computer Science and Artificial Intelligence Lab (CSAIL) have developed a wireless technology that promises to triple the speed of data transfers while also doubling signal range.

The researchers, led by Dina Katabi, an MIT professor of electrical engineering and computer science, recently demonstrated MegaMIMO 2.0, a new multiple-input, multiple-output (MIMO) technology that can coordinate several Wi-Fi routers at once, enabling the devices to triangulate data faster and more consistently. The new approach, joint multiuser beamforming (JMB), enables independent access points (APs) to beamform their signals and communicate with their clients on the same channel as if the APs were a single large MIMO transmitter.

In conventional wireless networks, multiple nearby transmitters cannot transmit simultaneously on the same frequency, since the signals would collide and become unreadable. MegaMIMO, however, is designed to enable multiple independent transmitters to transmit to multiple receivers at the same time and on the same frequency and still allow receivers to decode their signals.

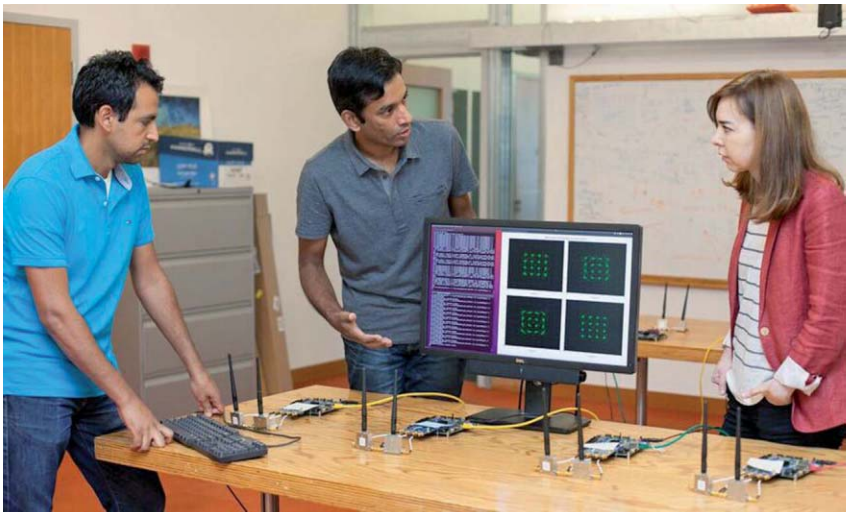
"Of course, the signals collide, which is unavoidable," says Hariharan Rahul, a former research team member and currently a visiting researcher with the project. "But MegaMIMO access points modify the transmitted signals so that at each receiver, after collision, only the desired signal to that receiver survives."

MegaMIMO promises a several-fold increase in wireless network throughput compared to existing wireless networks, says Rahul. "Further, MegaMIMO can do this simply by replacing the access points, and without requiring any hardware or software modifications to end user devices," he notes.

The key enabling technology behind JMB is a new low-overhead technique for synchronizing the phase of multiple transmitters in a distributed manner. The design allows a wireless LAN to scale its throughput by continually adding more APs on the same channel.

The researchers recently tested JMB with both software radio clients and off-the-shelf 802.11n cards in a deployment that simulated a densely congested conference room (Figure 3). Results from the ten-access point software-radio test bed showed a linear increase in network throughput with a median gain of 8.1 to 9.4×. The results also showed that JMB can provide throughput gains with standard, unmodified 802.11n cards.

MegaMIMO uses a variety of signal processing algorithms and techniques,



MIT

FIGURE 3. MegaMIMO 2.0 research team members doctoral student Ezzeldin Hamed, visiting researcher Hariharan Rahul and Prof. Dina Katab with a prototype of their technology, which promises to transfer wireless data over three times faster than existing systems.

including various properties of orthogonal frequency-division multiplexing signals, Fourier transforms and translation between time and frequency domains, as well as the efficient application of linear time-invariant filters to signals.

As they developed the technology, the researchers faced a challenge in coordinating time and phase synchronization. "The Wi-Fi transmitters have to be synchronized in time very tightly," Katabi says. Such synchronization must occur on a nanosecond scale. Also, since Wi-Fi signals are waves, two waves can combine to cancel each other out or, on the other hand, enforce each other. "If you are not careful about phase synchronization, the wave can combine to create the opposite of the intended effect," Katabi explains. MegaMIMO uses signal processing to enable access points to process signals in a synchronized manner, providing a lightweight, distributed approach that requires only minimal changes to the existing Wi-Fi wireless processing pipeline.

In real-world applications, MegaMIMO promises to dramatically improve throughput in the dense wireless networks—both Wi-Fi and cellular—commonly deployed in large, public places, such as sports stadiums, convention centers, hotels, airports, and shopping malls.

The researchers are now focused on scaling up the prototype system into larger deployments consisting of scores of Wi-Fi access points. "We have had interest from a variety of players in the wireless space, as well as end users that are currently facing challenges with dense wireless scenarios," Katabi says.

Author

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

SP

Digital Object Identifier 10.1109/MSP.2016.2647178

Hana Godrich, Arye Nehorai, Ali Tajer,
Maria Sabrina Greco, and Changshui Zhang

Special Article Series on Signal Processing Education via Hands-On and Design Projects

As professionals in the signal and information processing field, we build on the tools and theories we learned in our undergraduate studies, adding knowledge and skills over the years. For many of us, it has been a while since our undergraduate studies, yet we can probably recall some of these "Ah-ha" moments when a full meaning of some fundamental theory sank in. As educators and mentors, we seize on these moments and our own experiences to implement new successful teaching methods that advance effective learning and skill development.

In an era of unprecedented technology refresh rate, the challenge of providing a high-quality engineering experience is compounded by the required theoretical background, the increasing multidisciplinary applications, and a growing demand from the industry for engineers with "know-how" skills. Schools need to constantly assess and restructure the manner in which they prepare students to meet these growing demands from the engineering workforce. Educators are faced with the greater challenge of preparing their undergraduate students to deal with real-life engineering problems as early as possible in their education while not compromising on the required theoretical knowledge base. Engineering programs need to find an effective way to incorporate application aspects into the teaching of the fundamental concept. At the same time, they are required to develop teamworking skills, research and development experience, and innovative thinking, which is achievable through full-scale engineering design projects. For the latter, students need to collaborate on more complex engineering problems that integrate a larger set of tools and disciplines to solve and work under realistic constraints. The diversity in engineering applications that utilize signal and information processing opens many possibilities when it comes to the choice of experiments and projects that will keep students engaged in learning. The implementation of such practices becomes feasible with the availability of affordable hardware platforms that incorporate significant on-board computation capabilities alongside access to sensors, actuators, and open-source software tools.

There are few opportunities to share progresses and innovation made through undergraduate engineering design projects. An overview of the state-of-the-art methods used in providing students with practical engineering education is of high interest to educators, researchers, and professionals. A discussion on what is done around the world to advance students' hands-on experience will provide valuable tools and practices to educators and will offer professionals in the industry with a clearer image of the efforts made to increase engineering skills during undergraduate studies.

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* (*SPM*) through a series of article contributions from around the world.

Schäck, Muma, and Zoubir's article, "Signal Processing Projects at Technische Universität Darmstadt," details year-by-year practices implemented throughout undergraduate and graduate studies to support students' hands-on experience. The curriculum builds up theoretical knowledge alongside laboratories and engineering projects that advance professional proficiency. Interdisciplinary aspects, laboratories infrastructure, and the role of competitions in this process are discussed. This overview offers the reader an insight into use practices, detailing their advantages and challenges.

Focusing on engineering projects and competitions, Zhuo, Ren, Jiang, and Zhang's, article, "Hands-On Learning Through Racing," on the National Collegiate Intelligent Model Car Competition in China, introduces an education-through-challenge approach. In an annual competition, participating teams need to design and build cars that will be racing against other teams. The students learn a multitude of engineering skills while developing teamwork capabilities and collaboration skills. The article

Digital Object Identifier 10.1109/MSP.2016.2620199
Date of publication: 11 January 2017

offers extensive details on the structure of these competitions and the skill sets developed through it, enabling the adoption of this competition-based approach by others.

A focus on communications-related practices is given in the article, "Teaching the Principles of Massive MIMO" by Larsson, Danev, Olofsson, and Sörman. This contribution details the development of a course targeting students' exposure to cutting-edge technology and emerging concepts. The course is designed around building system-level understanding and expanding the classical curriculum to integrate a project-like approach. A student perspective is given throughout the article along with lessons learned. It demonstrates the students' experience and how students' feedback has been used to further develop the course impact.

Complementing this issue

Complementing these three feature articles are two articles published in *SPM*'s "SP Education" columns in the July and November 2016 issues, which paved the way for us to introduce the readers to this effort of sharing best practices on hands-on training in signal processing. In *SPM*'s July issue, Simoni and Aburdene [1] shared their eight-year experience and lessons learned in developing application-oriented activities to help students better understand signal processing theory and connect the theory to real-world applications.

In the November 2016 "SP Education" column, Richter and Nehorai introduced the incorporation of undergraduate research projects as a key component in the Electrical and Systems Engineering program at Washington University, St. Louis, Missouri [2]. Thanks to the active involvement of signal processing faculty members, many of the successful projects were related to signal processing, and these experiences substantially boosted the undergraduate enrollment and retention rate and attracted students to pursue a career in engineering.

Undergraduate engineering design projects, commonly introduced in a students' junior and senior years, allow them to work on real-life problems while applying their acquired knowledge and creativity. Some of these projects provide an opportunity to work in collaboration with others on more complex tasks, training students to learn teamwork skills and project management. These collaborations frequently entail a multidisciplinary effort. Signal and information processing plays an important role in many of these engineering projects.

While there are some channels in which students can share and publish their engineering projects, there is a need for a more focused review on engineering projects that offer great opportunities for the implementation of signal and information processing techniques. With the rapid advancement in technology and platforms available for project development, there is high value is sharing the knowledge and results stemming from these efforts to advance the general community. An overview of practical educational tools, application challenges, and keys to successful implementation of these programs is of high interest to both academia and the industry.

To address this need, as part of this article series, *SPM* has opened a SigPort-based submission and archival platform for sharing students' projects contributions. This issue's "SP Education" column is the first to detail these highlighted projects. Through the SigPort repository, a number of undergraduate students and their advisors shared information on relevant engineering projects. Overall, the submitted projects had more than 400 downloads within a two-month period, showcasing the keen interest in the community for such information. Contributions from around the world cover diverse fields and projects reflecting signal and information processing opportunities and applications range.

It is encouraging to learn that *SPM* and its monthly eNewsletter will be working with the IEEE Signal Processing Society's Education Committee and SigPort Committee to continue accepting student project submissions and theses in the broad areas of signal and information processing to archive through the SigPort platform. Summaries of the projects selected from these submissions will be periodically highlighted in *Inside Signal Processing eNewsletter*; and, if space allows, some of these projects may be showcased in *SPM*'s "SP Education" column.

We hope that the introduction of this series of articles dedicated to signal and information processing in engineering projects will promote communication and discussion on undergraduate studies, capabilities development, and increase interest and involvement from the engineering community. We look forward to bringing you the next set of informative articles in upcoming issues of the magazine.

Guest editors

Hana Godrich (godrich@rutgers.edu) received the B.Sc. degree in electrical engineering from the Technion Israel Institute of Technology, Haifa, in 1987; the M.Sc. degree in electrical engineering from Ben-Gurion University, Beer-Sheva, Israel, in 1993; and the Ph.D. degrees in electrical engineering from the New Jersey Institute of Technology, Newark, in 2010. She is the undergraduate program director in the Electrical and Computer Engineering Department at Rutgers University, Piscataway, New Jersey. Her research interests include statistical signal processing with application to wireless sensor networks, communication, smart grid, and radar systems. She is a Senior Member of the IEEE.

Arye Nehorai (nehorai@ese.wustl.edu) received the B.Sc. and M.Sc. degrees from the Technion Israel Institute of Technology, Haifa, and the Ph.D. degree from Stanford University, California. He is the Eugene and Martha Lohman Professor of Electrical Engineering in the Preston M. Green Department of Electrical and Systems Engineering at Washington University in St. Louis. He was editor-in-chief of *IEEE Transactions on Signal Processing* from 2000 to 2002. From 2003 to 2005, he was the vice president, Publications, of the IEEE Signal Processing Society (SPS), the chair of the Publications Board, and a member of the SPS's Executive Committee. He received the 2006 IEEE SPS Technical Achievement Award and the 2010

IEEE SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

IEEE SPS Meritorious Service Award. He was elected Distinguished Lecturer of the IEEE SPS (2004–2005). He received several best paper awards in IEEE journals and conferences. He is a Fellow of the IEEE, the Royal Statistical Society, and American Association for the Advancement of Science.

Ali Tajer (tajer@ecse.rpi.edu) received the M.A. degree in statistics and Ph.D. degree in electrical engineering from Columbia University, New York, in 2007 and 2010, respectively. He is currently an assistant professor of electrical, computer, and systems engineering at Rensselaer Polytechnic Institute, New York. His research interests include mathematical statistics and network information theory with applications in wireless communications and power grids. He serves as an editor for *IEEE Transactions on Communications* and *IEEE Transactions on Smart Grid* and as the guest editor-in-chief for *IEEE Transactions on Smart Grid* (special issue on theory of complex systems with applications to smart grid operations). He received the NSF CAREER Award in 2016. He is a Senior Member of the IEEE.

Maria Sabrina Greco (m.greco@iet.unipi.it) graduated with a degree in electronic engineering in 1993 and received the Ph.D. degree in telecommunication engineering in 1998, from the University of Pisa, Italy, where she has been an associate professor since December 2011. She is an associate editor of *IET Proceedings–Sonar, Radar, and Navigation*, editor-in-chief of *IEEE Aerospace and Electronic Systems Magazine*, member of the editorial board of the Springer *Journal of Advances in Signal Processing*, and senior editorial board member of *IEEE Journal on Selected Topics of Signal Processing*. Her general interests are in the areas of statistical signal processing, estimation, and detection theory. She coauthored many book chapters and more than 170 journal and conference papers. She is a Fellow of the IEEE.

Changshui Zhang (zcs@tsinghua.edu.cn) received the B.S. degree in mathematics from Peking University, Beijing, China, in 1986 and the Ph.D. degree from the Department of Automation, Tsinghua University, Beijing, in 1992, where he is currently a professor. He has authored more than 200 papers published in journals and conferences. His current research interests include machine learning, artificial intelligence, image processing, pattern recognition, and evolutionary computation. He is also an associate editor of *Pattern Recognition*.

References

[1] M. Simoni and M. Aburdene "Lessons learned from implementing application-oriented hands-on activities for continuous-time signal processing courses," *IEEE Signal Process. Mag.*, vol. 33, no. 4, pp. 84–89, July 2016.

[2] E. Richter and A. Arye Nehorai, "Enriching the undergraduate program with research projects," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 123–127, Nov. 2016.

SP

Digital Object Identifier 10.1109/MSP.2016.2643718

SIGNAL PROCESSING EDUCATION
VIA HANDS-ON AND DESIGN PROJECTS

Tim Schäck, Michael Muma, and Abdelhak M. Zoubir

Signal Processing Projects at Technische Universität Darmstadt

How to engage undergraduate students in signal processing practice

MAIN IMAGE ©ISTOCKPHOTO.COM/SKYNESHER; SCREEN IMAGE ©ISTOCKPHOTO.COM/ALENGO

Digital Object Identifier 10.1109/MSP.2016.2619218
Date of publication: 11 January 2017

This article is meant to share our experience on integrating signal processing hands-on opportunities into formal engineering education at Technische Universität (TU) Darmstadt, Germany. As many universities face the challenge of how best to provide hands-on experience to undergraduate students, we hope to inspire our colleagues and perhaps trigger new hands-on projects by sharing our insights. At TU Darmstadt, we believe that it is essential to provide undergraduate students with hands-on signal processing opportunities right from the starting point of their studies through graduation.

Introduction

Hands-on education in signal processing has a long-standing tradition (e.g., [1]–[5]), and its importance, given the complexity of today's engineering problems, is undisputed. At TU, we hope that we can—in one way or another—inspire some of our colleagues who are involved in educating the next generation of signal processing researchers and practitioners.

We will briefly explain the format of the projects and highlight some important challenges in the implementation as well as successful strategies and pitfalls that we encountered. The time line of the curriculum, as shown in Figure 1, serves as a structure to present material in an ordered fashion. However, all sections can be read independently. We also illustrate how we utilize student competitions, such as the IEEE Signal Processing Cup, to stimulate innovation and collaboration between graduate and undergraduate students. Special attention is given to the many possibilities that collaborations with industry partners offer for students. The involvement of students in interdisciplinary research, which has a long-standing tradition at TU Darmstadt, is illustrated by the example of a cooperation between the signal processing group and the psychology group.

Laboratories are central to our hands-on education for freshmen to senior-year students. By promoting and extending our labs, we increase the exposure to state-of-the-art research and advanced equipment to our undergraduate students. For this reason, a separate section is devoted to presenting our diverse signal processing laboratories and the opportunities they offer our students. Some of the laboratories are purely educative, while others additionally provide the advanced students with hands-on research opportunities. The real-data experiments include fascinating topics, such as real-time audio signal processing, multiantenna receive beamforming, biomedical signal processing, geolocation, and tracking, to mention a few. Our newest laboratory considers the cutting-edge research topic of bioinspired communication. Here, students can generate their own data by performing single-cell experiments involving fluorescence microscopy and microfluidics. Based on their own data sets, the students can develop advanced signal and image-processing algorithms, e.g., for segmentation and tracking of single cells. We conclude with some practical general remarks on some fundamental aspects that we have found to be important for successful design projects. Short interviews as well as photos and figures are used to make the article an enjoyable and informative read after a hard day inside signal processing.

At TU Darmstadt, we believe that it is essential to provide undergraduate students with hands-on signal processing opportunities right from the starting point of their studies through graduation.

Signal processing within the curriculum of electrical engineering and information technology at TU Darmstadt

In Germany, the format of undergraduate education in electrical engineering and information technology (ETiT) was traditionally the diploma degree with a duration of five years. More recently, due to changes relating to the Bologna Declaration, a more internationally acknowledged system has been installed. It includes a three-year bachelor's degree followed by a two-year master's degree as new course structures. In this section, we describe how hands-on signal processing projects are integrated into the curriculum of ETiT at TU Darmstadt. At the end of each project, we list the positive aspects with ⊕, pitfalls with ⊖, and additional hints with ⊙.

First year

At TU Darmstadt, we believe that it is essential to provide undergraduate students with hands-on opportunities right from the starting point of their studies. During this phase, freshmen are especially motivated and highly curious. The lack of fundamental knowledge in engineering and science is

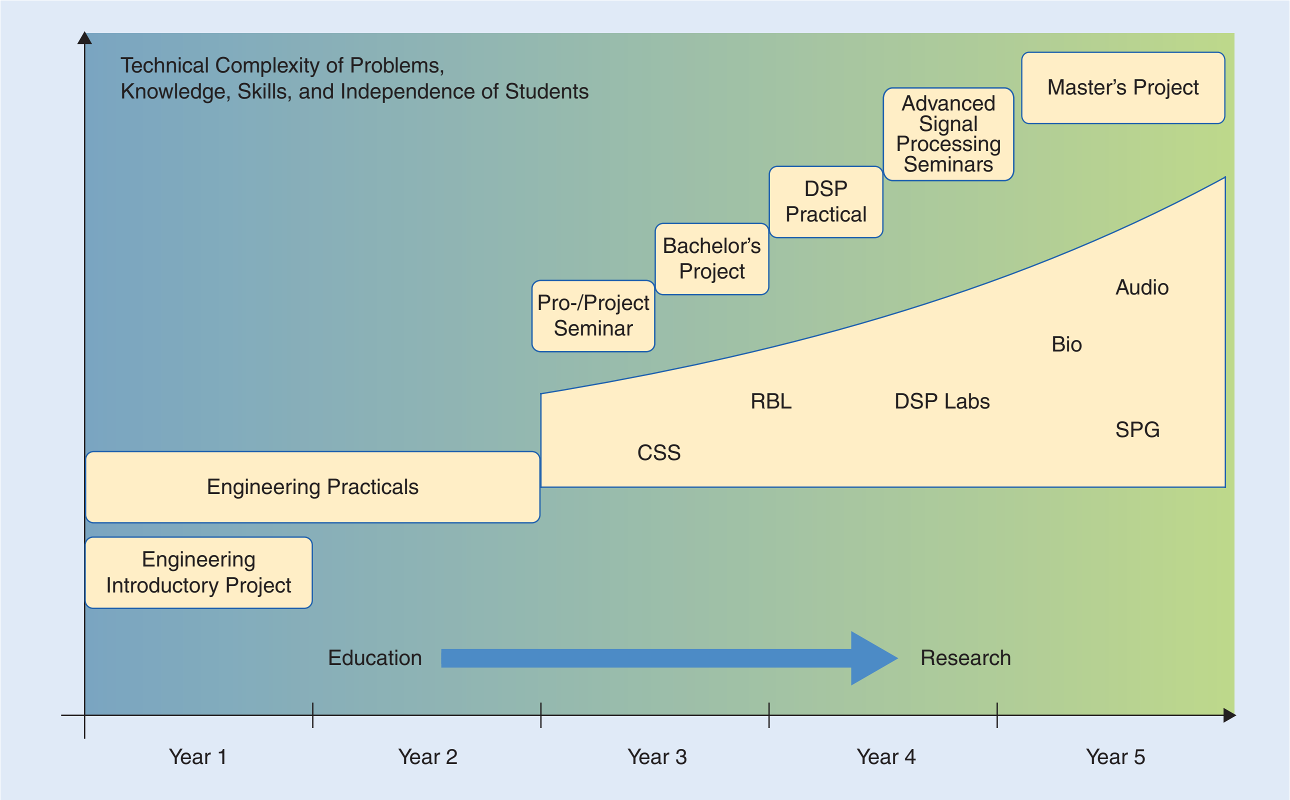


FIGURE 1. An overview of the hands-on activities in signal processing within the curriculum of electrical engineering and information technology at TU Darmstadt. We offer a variety of Digital Signal Processing (DSP) Labs: the Communication and Sensor Systems (CSS) Lab, the Receive Beamforming Lab (RBL), the Bioinspired Communication Systems (Bio) Lab, the Advanced Real-Time Audio Processing (Audio) Lab, and the Signal Processing Group (SPG) Lab.

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

often compensated by common sense coupled with creativity. Taking all of this into account, the Department of Electrical Engineering and Information Technology created an introductory project for freshmen in 2007.

Engineering Introductory Project

In the Engineering Introductory Project, interdisciplinary groups of approximately ten students work together on a technical solution to a timely, practical, complex, and socially relevant problem (Figure 2). The project takes place when students are only one to two months into their first semester, and it is a welcome contrast to the fundamental coursework that is usually offered in this early phase of the curriculum. The overall number of participating students is approximately 500.

Through the interdisciplinary exchange, students are given the opportunity not only to improve their technical skills but also to develop soft skills, such as teamwork and self-organization.

The focus of the Engineering Introductory Project does not rely on technical details. Rather, the first insights into today's engineering work in an interdisciplinary environment is provided. The project also gives the opportunity to make friends with classmates and to establish a first contact with the research associates (RAs) who serve as supervisors. RAs in Germany are appointed to assist professors with teaching, research projects, and, at the same time, to pursue a doctoral degree (Dr.-Ing.). This path excludes formal classes, thus, the term *research associates*. The peers in the group learn together and make practical experiences, which creates a pleasant and inspiring environment, an inevitable requirement for creativity.

The topic of the Engineering Introductory Project is not announced before the start of the project. Each team has one week to jointly develop an innovative solution. The team's final results are formally presented in front of a jury that is composed of professors. Starting in 2007, the topics of the Engineering Introductory Projects were as diverse as developing a power supply package for outdoor holidays (2008), to a cookie-baking machine (2010), to contributing to future living (2013).

FIGURE 2. An interdisciplinary group of freshmen working together on their Engineering Introductory Project. (Photo courtesy of Paul Glogowski/TU Darmstadt.)

During the project, two RAs serve as advisers for each group. On one hand, the soft skills adviser, i.e., an RA from the Department of Humanities, assists in creating an encouraging group dynamic and helps the team in reflecting their teamwork and interactions. On the other hand, the technical adviser, i.e., an RA of the Department of Electrical Engineering and Information Technology, answers questions regarding the technical aspects and encourages the group to use engineering tools. Members of the SPG participate as technical advisers to guide and motivate the freshmen with a special focus on signal processing.

Through the interdisciplinary exchange, students are given the opportunity not only to improve their technical skills but also to develop soft skills, such as teamwork and self-organization. In this way, the students get an impression of what awaits them later as professional engineers. Also, students begin to network at an early stage, even between different disciplines.

This project not only fosters didactic and technical learning in a team with other students of different interests but the freshmen also receive expert guidance from the professors of ETiT in dedicated consultation hours. Here, they are given the opportunity to discuss their ideas and ask questions related to their project with a specialist in this particular area. For example, Prof. Abdelhak Zoubir offers a consultation hour in which freshmen ask questions about challenges related to the field of signal processing that they have identified within their project.

One challenge in implementing this design project is the choice of an appealing and trendsetting topic. Its technical complexity must be adjusted to the students' knowledge and the given time frame. To test and evaluate possible solutions and to discover potential pitfalls throughout the project, the soft skills and technical advisers simulate the Engineering Introductory Project task beforehand within a period of three days. Their experience flows back into the project description and enhances the quality of the design project.

- ⊕ Freshmen practice working independently in interdisciplinary teams
- ⊕ a first exposure to signal processing problems
- ⊖ limited prior knowledge is assumed
- ⊙ the choice of topic is essential.

Second year

Before students in electrical engineering can be exposed to real-world problems in signal processing, they have to study the fundamentals of signal processing. In the second year, our students learn the basic concepts of signal processing by taking the course, "Deterministic Signals and Systems" and "Fundamentals of Signal Processing."

Fundamentals of the signal processing unit

The course, "Deterministic Signals and Systems," teaches undergraduate students the principles of deterministic signals and system theory. It starts with the Fourier series and the Fourier transform, treats linear time-invariant systems and convolution, and gives an overview of other signal transformations, such as the Laplace transform, the z-transform, and the discrete-time Fourier transform. Students apply these transformations to solve tasks related to physical problems that are modeled by linear differential equations.

Subsequently, the course, "Fundamentals of Signal Processing," covers basic concepts in signal processing, such as random variables, stochastic processes, random signals and linear time-invariant systems, optimal linear systems such as the matched filter and the Wiener filter, and the method of least-squares. Practical experiments with real-world data are shown in the lecture to help the students grasp the basic ideas in an intuitive fashion. The aim of the lecture is, furthermore, to serve as an introductory course for more advanced lectures in DSP, adaptive filtering, communications, and control theory.

As for many other fundamental courses in ETiT at TU Darmstadt, the lectures are complemented by tutorials that are held by selected undergraduate teaching assistants (UTAs) of higher semesters. The advantage of having students do the teaching instead of RAs or professors is that higher-semester students are aware of the difficulties from their own recent experience and can adequately support the more junior students. By recruiting good students as UTAs, signal processing can also be effectively advertised to more junior students. The UTAs who run the tutorials strengthen their knowledge in signal processing and integrate more into our research group. They usually conduct their bachelor's or master's thesis project with us and participate in other research projects or competitions.

- ⊕ Practical experiments with real-world data in the lectures help students grasp basic ideas in an intuitive way
- ⊖ the involvement of UTAs as a means of integration
- ⊙ the basic knowledge of signal processing is still missing at this point.

Third year

In the third year, we offer a variety of hands-on opportunities ranging from practical signal processing experiments in laboratories to small-scale research projects (Proseminar/Project Seminar) and the bachelor's thesis project. Students also can find their first exposure to interdisciplinary research projects in the Forum for Interdisciplinary Research Project (see "The Forum for Interdisciplinary Research Project"). In this section, we briefly explain the format of the third-year projects, highlight the most important challenges in their implementation, and discuss our own successes and pitfalls.

Communication and sensor systems laboratory

This practical consists of eight fundamental hands-on experiments from the field of communication engineering and signal processing: 1) the localization of acoustical sources, 2) digital modulation, 3) multiple-input, multiple-output communication, 4) software-defined radio, 5) parasitic effects in passive radio-frequency (RF) devices, 6) polarization of light, 7) RF field-effect transistor amplifier, and 8) the fields and impedance of antennas. The students are guided to acquaint themselves with each topic and are required to write reports about the conducted experiments.

To illustrate the hands-on signal processing opportunities offered to the students, consider the localization of an acoustical sources experiment. Here, students are given the opportunity to localize acoustical sources in our laboratory. To this end, the third-year students estimate the time-differences of arrival and angles of arrival (AOA) using correlation and generalized correlation functions. For the final audio source localization, the students fuse multiple AOA measurements from distributed microphone arrays, as shown in Figure 3. The eight microphones are divided into pairs that are mounted on the four walls of the laboratory (see Figure 4). The positions of the microphone arrays are given by $p_1, \ldots, p_4$, whereas the sound source is located at some unknown position in the center of the room. The measurements are recorded with eight standard Behringer B-5 condenser microphones and a Behringer Ultragain Pro-8 digital device, which is an eight-channel analog/digital and digital/analog converter. All calculations are performed in MATLAB.

In this laboratory, third-year students experience their first hands-on experiment. They can apply the freshly learned fundamentals of signal processing and have to submit their results in a written report. This lab also fosters the ability to work in teams. There exist some pitfalls regarding teamwork for hands-on signal processing. First, it might occur that teams do not distribute the work load evenly between the team members. Second, the team members sometimes split up the work such that only some members run the hands-on experiments while others write the report. In such cases, the laboratory adviser needs to remind the students to participate in the hands-on experiments at every stage.

- ⊕ The first hands-on experience in a DSP laboratory
- ⊕ the fundamentals of signal processing are practiced using real experiments
- ⊕ fosters the ability to work in teams
- ⊖ sometimes weaker team members are less active during the experiments
- ⊙ workload should be evenly distributed among team members.

As preparation for larger undergraduate research projects, such as the bachelor's thesis project, TU Darmstadt has introduced the Proseminar and Project Seminar. These are described in the following sections.

Proseminar

Scientific work always starts with understanding the state of the art. The students are expected to be informed about the research that has already been conducted in the field of interest, first, before reimplementing successful methods or even examining

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

The Forum for Interdisciplinary Research Project

Clearly, the most obvious way to provide students with hands-on experience in signal processing is to let them solve real-world problems. Interdisciplinarity naturally comes into play when dealing with hands-on research projects that solve real-world problems, and it requires different types of expertise to be combined. TU Darmstadt has a long-standing culture of cooperation across department boundaries. In this way, we supplement the classical department structure through flexible cooperation forms for research and teaching. The Forum for Interdisciplinary Research (FiF) builds on the successful interdisciplinary work at TU Darmstadt. The FiF was founded as a result of a senate decision in December 2008.

For example, in 2012, FiF funded a research cooperation between Prof. Abdelhak M. Zoubir (signal processing) and Prof. Augustin Kelava (psychology). The joint research topic was to investigate the synchronization of physiological signals in emotional situations. Emotion-eliciting situations are accompanied by changes of multiple variables associated with subjective, physiological, and behavioral responses. The quantification of the overall simultaneous synchrony of psychophysiological reactions plays a major role in emotion theories and has received increased attention in recent years. From a psychometric perspective, the reactions represent multivariate nonstationary intraindividual time series.

Undergraduate signal processing and psychology students, supervised by research associates (RAs) and professors, cooperated to robustly determine the synchrony of psychophysiological reactions. Within the FiF project, bachelor's and master's thesis projects were undertaken in both departments. The cooperation quickly revealed that processing psychophysiological data is challenging. Often motion artifacts affect physiological time series. However, the data also offers the hands-on possibility to study many fundamental concepts, such as filter design, time-frequency and wavelet analysis, adaptive filtering, robust statistics for dependent data, parameter and signal estimation, detection, and classification.

FIGURE S1. An undergraduate student at TU Darmstadt records physiological signals while watching an emotion-inducing video clip.

One of the signal processing projects, which we conducted during a master's thesis project within the FiF, focused on motion artifact removal in electrocardiographic (ECG) signals. The developed method had to provide satisfactory results for a large range of data and also be computationally efficient. Further, it had to be programmed in a way such that psychology students who did not have any signal processing background would be able to use it. All of these requirements together provided a realistic hands-on framework to conduct a master's thesis project, which resulted in a publication at the European Signal Processing Conference 2012 [7].

We observed that the interdisciplinarity of the FiF project, as well as the feeling of being able to solve real problems, created a unique team spirit among the students. Even today, signal processing students use the databases that were established to develop and evaluate new methods. Also, psychologists use the signal processing methods to study the synchronization of psychophysiological signals in the body in many different emotion-eliciting situations (Figure S1). New challenges and opportunities arise from the possibility to integrate wireless body-worn sensors into the psychological experiments. This allows the psychologists to undertake more realistic experiments and provides the signal processing students with new and even more challenging data sets.

- ⊕ Psychophysiological data is challenging and requires the study of fundamental concepts
- ⊕ databases, once established, can be reused by other signal processing students
- ⊕ the interdisciplinary nature of the project requires explaining fundamental concepts without using equations
- ⊕ excellent results could be obtained, and publications as well as bachelor's and master's thesis projects were produced
- ⊖ it takes a long time until students from both research fields speak the same language
- ⊖ student projects are of a short duration and tend to conclude when a student has reached his/her peak in productivity
- ⊙ documentation of the methods and datasets is essential in interdisciplinary research.

new approaches. This literature survey is the subject of the Proseminar that, as a first step toward scientific work, lays the foundation for performing hands-on research projects later on. It takes about four weeks to accomplish the Proseminar.

During the Proseminar, students read books, papers, or complementary work on a given subject in signal processing. The topics are not limited to fundamentals but may also cover very recent papers and approaches. To finalize the Proseminar, the third-year students summarize their results in a written report that is checked by the supervising RA.

Through the intensive literature survey, the students learn to understand, analyze, and summarize scientific papers. However, the papers must be carefully selected. They are usually aligned with the doctoral research of the supervising RA. Another challenge is to define a project such that it can be completed within four weeks while also considering the students' prior knowledge.

Project Seminar

After having reviewed the literature in the Proseminar, students investigate and solve a specific signal processing problem by reproducing an already existing approach. They typically reimplement an algorithm from a published paper and try to reproduce its results. During the project, students are allowed to suggest their own modifications and extend the methods to improve the results. Such creative contributions yield a bonus when it comes to grading. However, their own contributions are not required and are primarily the aim of the subsequent bachelor's thesis project. The reproducibility of the publications is essential, and the paper content and quality must be checked by the supervising RA beforehand.

In addition to the reimplementation, students search for and analyze scientific reference publications and, in the end, summarize the obtained results and their conclusions in a written report. The outcomes of the Project Seminar are defended in front of the research group and students in an oral presentation. The duration of the Project Seminar is about two months.

For the students, the Project Seminar is one of the first opportunities to develop skills in MATLAB, Latex, and BibTeX, which are necessary tools for scientific work in signal processing and the basis for subsequent hands-on projects. The aim of the Project Seminar is to practice applying methods of signal processing to practical problems and to gain some knowledge in a particular research area, which can be built upon in later projects. As for the Proseminar, the challenge is to define the project such that it can be completed on time. Often, the workload for the RA who supervises the projects is high, as regular meetings with students are essential to ensure a high quality of work. Also, the corrections of the report often include tedious linguistic corrections, since the students are not yet acquainted with scientific English.

- ⊕ Develop skills in MATLAB, Latex, and BibTeX
- ⊕ apply signal processing methods to practical problems
- ⊕ deepen knowledge in a particular field
- ⊖ the workload for the RA who supervises the projects is high in relation to the outcome.

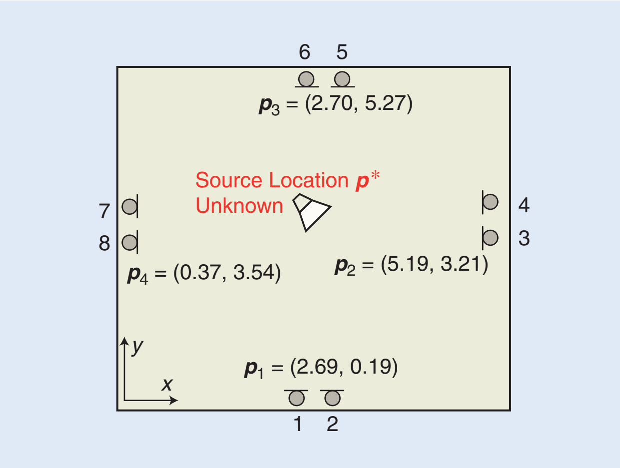


FIGURE 3. The setup of the Localization of Acoustic Sources experiment that is part of the Communication and Sensor Systems Laboratory offered to third-year students.

FIGURE 4. A microphone pair of the Localization of Acoustic Sources experiment that is attached to one of the walls of the Signal Processing Laboratory.

Bachelor's thesis project

The third year concludes with the bachelor's thesis project, which usually builds upon the Proseminar and Project Seminar. The bachelor's thesis project is designed to last three months. The length of the bachelor thesis is typically 40–80 pages. Approximately 100 students in ETiT finish their bachelor thesis each year. Similar to the Project Seminar, the students have to give a 20-minute presentation and defend their work in front of an audience. The bachelor's thesis project offers the possibility for students to be creative and to develop new ideas and algorithms. Students can either explore new ways of solving a specific problem, compare different methods and their performance, or improve an existing approach by extending or enhancing particular aspects. In general, the bachelor's thesis project is based on research questions provided by the RA. In our group, we offer hands-on experiments using, e.g., biomedical, audio, or ultrasound data, which can be acquired by the students in our signal processing lab. Outstanding bachelor's thesis projects can lead to publications and visits to conferences such as the IEEE Workshop on Statistical Signal Processing, the European Signal Processing Conference (EUSIPCO), or the International Conference on Acoustics, Speech, and Signal Processing (ICASSP).

One of our undergraduate students who was published at ICASSP'14 was Jack Dagdagan (see Figure 5). In his paper [6], he developed a robust method for testing stationarity in the presence of outliers. Jack recollects, "In my bachelor's thesis project, I evaluated my algorithm first with simulated data and assumed a certain outlier model. I was not sure whether the method would still perform reasonably well when using real data. So I was very excited when I started recording real data. I noticed that the computational complexity was much higher with real data, since the simulated data had a sample size of only 1,024 and I recorded some seconds of audio with a 48,000-Hz sample rate. I realized that the outliers were completely different to the outliers in the simulations. But I was very happy that the performance of my method was still very good. If you develop an algorithm and evaluate with simulations, you could optimize by tweaking your parameters such that your algorithm fits the model perfectly. However, if your model is very special, your method won't work in reality. But if you test your algorithm on real data and still achieve good results, it shows that either your model was very diverse or that your algorithm runs very well independently of the model you are using."

Another lesson that Jack learned from his hands-on experience is the difference between theory and practice. "In theory, you learn the definitions of stationarity, such as wide-sense stationarity (WSS)," he says. "However, in practice, you realize that you can never have perfect WSS. Thus, you need to set a threshold above which you determine the signal not to be stationary anymore. Before my hands-on experience in this bachelor's thesis project, I would not have thought of stationarity in this way."

- ⊕ Outstanding bachelor's thesis projects that use real-world data can lead to publications and conference visits

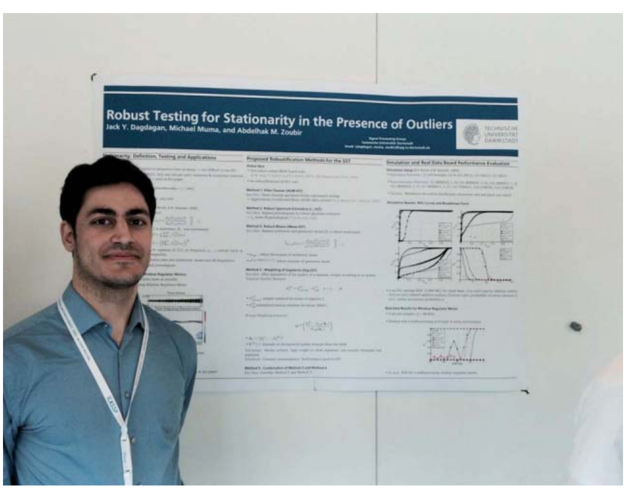


FIGURE 5. Jack Dagdagan, an undergraduate student at TU Darmstadt, presenting the results of his bachelor's thesis project at ICASSP'14 in Florence, Italy.

- ⊕ students can build upon the work they did in the Proseminar and Project Seminar
- ⊙ acquiring real data and working with it must be well planned, otherwise it would exceed the three-months nominal length
- ⊙ all of the aforementioned projects, i.e., Proseminar, Project Seminar, and the bachelor's thesis project, are graded after the seminar by the professor.

Fourth and fifth year

The following programs and activities are part of the master's program of ETiT, which forms the final two-year stage of the undergraduate education at TU Darmstadt.

In their fourth and fifth years, signal processing students at TU Darmstadt are ready to tackle some more challenging and realistic problems.

DSP Practical

Fourth-year students can attend the DSP Practical, either in parallel to or after the course "Digital Signal Processing." It offers the chance to further familiarize onself with signal processing programs, such as MATLAB, and put theory into practice. Students participating in this lab are able to apply the concepts from the lectures. This covers mainly the design of finite impulse response and infinite impulse response filters as well as parametric and nonparametric spectrum estimation; examples of the latter are shown in Figure 6. Real-world signals, such as speech and audio signals, touch-tone telephone dialing signals, temperature recordings, or biomedical measurements, are either provided to or recorded by the students. UTAs help to supervise the undergraduate students during the experiments. For example, the biomedical experiment, where students record each others electrocardiogram (ECG) signal and perform spectrum estimation, was designed with the help of a UTA. The experiments are conducted in the SPG Lab, which is described in the "Laboratories" section.

The DSP Practical is composed of eight practicals and two real-data acquisition sessions. Approximately ten groups of two to three students work together to solve signal processing tasks. As an introductory part for every experiment, students receive handouts with the underlying theory and some preparatory questions. The students' understanding of the theory is checked by the supervisor at the beginning of each experiment. In this way, we ensure the students' adequate preparation for the practicals. Furthermore, for every experiment, each group writes a report in which they wrap-up their results, answer all questions, and include plots and code from the experiment. At the end of the semester, a final exam is held.

- ⊕ Students are able to apply concepts learned in the DSP course using real-world data
- ⊕ students can further familiarize themselves with signal processing tools, such as MATLAB
- ⊕ students can acquire their own measurements
- ⊖ the time slot per experiment is tight
- ⊖ tasks are explicitly predefined, and the time for trial and error is very limited.

Advanced seminars in signal processing

In their fourth and fifth years, signal processing students at TU Darmstadt are ready to tackle some more challenging and realistic problems. At this time, they have acquired sufficient fundamental knowledge, and they are also used for reading and reproducing results from papers. The SPG offers an advanced seminar in which master's students tackle small-scale real-world problems to prepare students for their master's thesis project, as well as to provide the opportunity to get to know the RAs and practice working in teams. The seminar, "Advanced Topics in Statistical Signal Processing," is aimed at students who have an interest in signal processing and a desire to extend their knowledge of signal processing in preparation for future project work, e.g., their master's thesis project and their working careers. The seminar consists of a short series of lectures (i.e., four to five), followed by student group projects (six to eight weeks), a presentation of the achieved results, and a final exam. Usually, up to 20 students participate in the seminars. The topics of the lectures and the student projects are different every year. The RAs are free to propose students' projects, and the students make their choice based on their own interest. In this way, both the students and their supervisors are highly motivated.

Students often use the SPG Lab to investigate topics such as direction-of-arrival estimation or localization of sound sources in impulsive noise environments. In many cases, students become creative in the experiments. For example, one group used their mobile phones to play sound files (both signal and noise) in combination with a miniature train to create audio sources that moved on a fixed trajectory.

- ⊕ Deepen the knowledge in signal processing
- ⊕ preparation for final year master's thesis project
- ⊕ groups work on a common research topic and present results
- ⊕ contact with RAs is intensified
- ⊖ time is often too short for the students to gain a deep insight.

Final year master's thesis project

The four-semester master's program in ETiT consists of compulsory core courses, compulsory elective courses, and elective courses as well as the master's thesis project. In their master's thesis project, the students work independently for a duration of six months on a scientific project under the supervision of one of the RAs. The research topics are larger and more complex compared to the bachelor's thesis project.

For many students, the master's thesis project offers the possibility to conduct research on real-world data. From our experience, the best results are obtained when the students are involved in collecting their own data. In this way, they acquire hands-on contextual information and can better understand the data, e.g., in terms of the signal quality, the measurement principle or the assumptions on the noise distributions. Further, when the master's thesis project solves a real-world problem, the students take more care that the developed algorithms are designed in accordance with practical requirements. The latter include, e.g., computational efficiency of the algorithm, communicational load between sensors, memory restrictions, or real-time requirements. In our experience, the students enjoy incorporating such realistic requirements into their algorithm design.

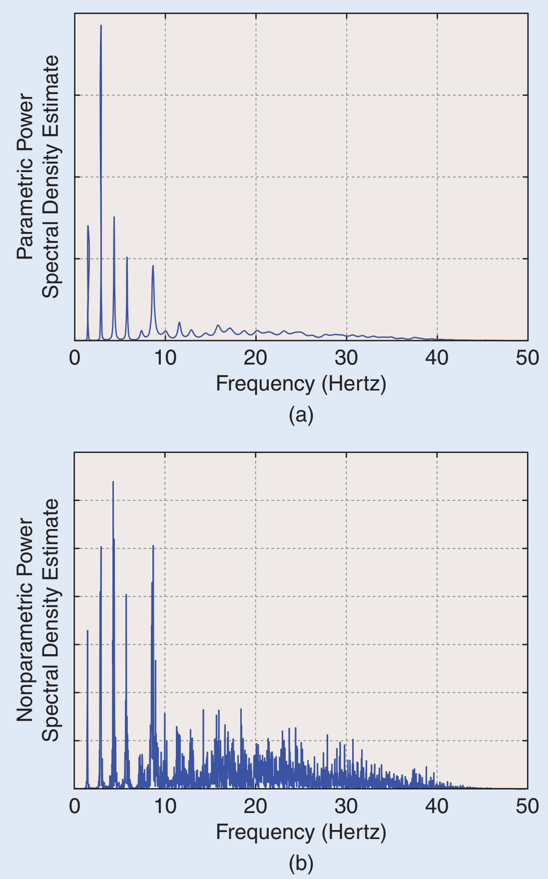


FIGURE 6. Examples of (a) a parametric power spectral density estimate and (b) a nonparametric power spectral density estimate. Both estimates were computed during the DSP Practical with data collected by the undergraduate students.

It is not uncommon that the projects are carried out in cooperation with a selected industry partner or a research center. In case of such cooperation, it must be emphasized that we take special care to make sure that the master's students perform real-world signal processing research tasks under the supervision of a qualified supervisor. For this reason, cooperations for master's thesis projects are only possible with trusted institutions with which a solid research partnership has been established.

At TU Darmstadt, the students have the unique possibility to choose between a broad range of hands-on topics at national and international locations. Examples of past SPG master's thesis projects of Prof. Zoubir in cooperation with research or industry partners are as diverse as

- signal processing for photometric glucose measurement in hand-held devices at a German pharmaceuticals and diagnostics company

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

- room shape estimation in through-the-wall radar imaging in cooperation with the Center for Advanced Communications, Villanova University, United States
- directional integration of wirelessly transmitted audio signals into hearing aid processing, and many more.

Clearly, each of these applications has its own challenges.

"When I was a master's student at TU Darmstadt," remembers Dr.-Ing. Michael Muma, now a postdoctoral researcher with the SPG, "I was walking through the corridor and saw a notice that immediately caught my attention. The notice literally opened my eyes to an exciting application of signal processing: the research of human vision. Doing my master's thesis project in the Contact Lens and Visual Optics Laboratory (CLVOL), which is part of the School of Optometry and Vision Science at Queensland University of Technology, Brisbane, Australia, was an amazing experience. The CLVOL has a sophisticated range of unique measurement and analysis techniques. These include methods to investigate the shape of the cornea, the optical characteristics of the eye, visual performance of contact lenses, and the biometric properties of the eye. After all the coursework, I really wanted to test my skills in a practical environment, and I also wanted to see how research was organized outside the university. While I was at the CLVOL, I took all the measurements myself. This involved the synchronous measurement of the eye's wavefront aberrations, cardiac function, blood pulse and respiration signals. We were among the first ones to analyze the role of cardiopulmonary signals in the dynamics of wavefront aberrations [8]. My master's thesis project was under the supervision of Dr. Robert Iskander, who is now a professor at Wroclaw University of Technology." A schematic sketch of the relation between the eye's wavefront aberration and the ECG, the BP, and respiration signals is shown in Figure 7.

"Prof. Iskander has his own teaching philosophy [9], and, after the master's thesis project, it was clear to me that I wanted to continue with signal processing research. Working with real data in a team consisting of engineers and eye researchers allowed me to grasp the importance of signal processing. Today's measurement devices and data analysis are too complex to be handled without a rigorous understanding of signals and systems theory. At the SPG, it is important that students are given the opportunity to work hands-on from the start. I am very grateful that I could do my master's thesis project in such an inspiring environment full of high-tech custom equipment," says Dr. Muma.

For the development of DSP algorithms that show good and robust performance in real product applications, real-time tests with typical signals in realistic environments are essential.

- ⊕ Students are ready to undertake larger projects and to work independently
- ⊕ cooperation with selected industry and research partners provides hands-on experiences
- ⊕ outstanding projects that use real-world data can lead to publications and conference visits
- ⊕ recruitment of RAs
- ⊖ master's students often stop at the height of their productivity.

Laboratories

Central to our hands-on education, from freshmen to senior-year students, are the signal processing laboratories at TU Darmstadt. This section presents our diverse signal processing laboratories and the opportunities they offer for our students.

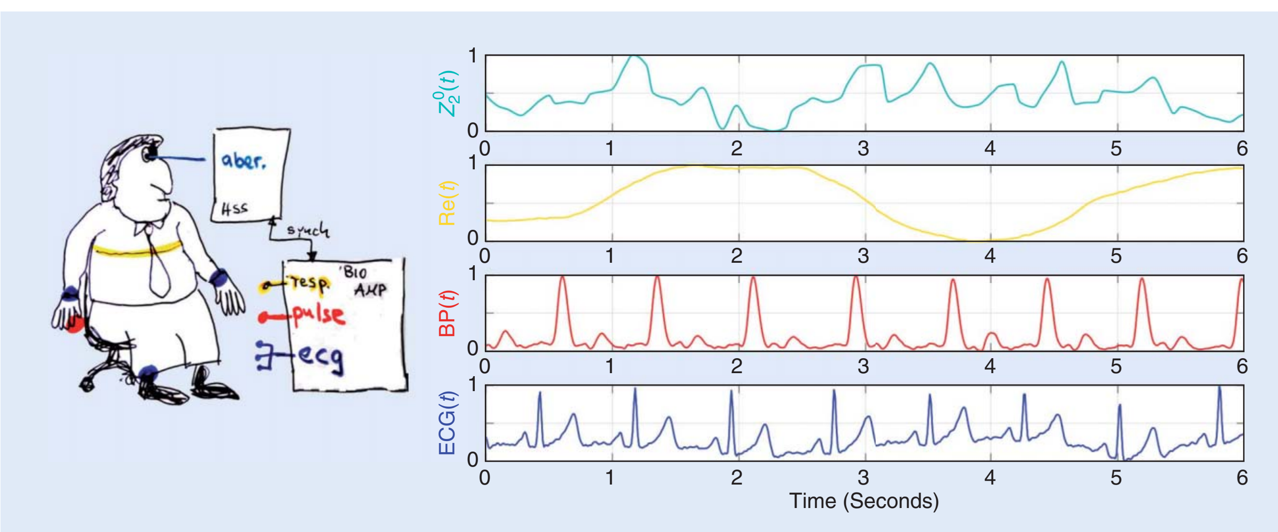


FIGURE 7. An example of a hands-on master's thesis project undertaken in cooperation with the CLVOL at Queensland University of Technology, Brisbane, Australia. The picture shows a schematic sketch of the defocus component of an eye's wavefront aberration $Z_2^0(t)$, the respiration signal Re(t), the blood pulse BP(t), and the ECG(t).

As detailed next, some of the laboratories are mainly educative in nature, while others additionally provide the advanced students with hands-on research opportunities and expose them to state-of-the art research and advanced equipment.

The SPG Lab

In our group, we offer hands-on experience to the undergraduate students in the SPG Lab, consisting of a biomedical sensor lab, a basic audio signal processing lab, and a synthetic aperture sonar lab. It is mainly funded by the so-called resources for quality assurance of study and teaching (QSL)—essentially, a fund for enhancing hands-on experiences—in teaching. In the biomedical sensor lab, a variety of sensors offer the opportunity to acquire own measurements, such as ECG, photoplethysmogram (PPG), and blood pressure. The data is recorded using ADInstrument devices that are originally designed for research and teaching at universities.

- ⊕ Students can use the SPG Lab for their bachelor's or master's thesis projects
- ⊕ students can collect their measurements
- ⊕ undergraduate students can jointly carry out research with RAs
- ⊕ even patented technologies have been developed in this lab
- ⊙ keeping the lab up-to-date and providing GUIs and help to students is time consuming.

Advanced real-time Audio Processing Lab

For the development of DSP algorithms that show good and robust performance in real product applications, real-time tests with typical signals in realistic environments are essential. Those real-time tests show a variety of natural setups that cannot be covered by data simulations. Prof. Henning Puder's research group, Adaptive Systems for Speech and Audio Signal Processing, provides such a system for the development of audio processing algorithms in students' projects. The core component is a real-time DSP system, Speedgoat [10], with 12 analog audio input and eight output signals (see Figure 8). The signals are processed with low latency (<1 millisecond). Algorithms can be implemented in high-level programming languages such as MATLAB/Simulink. To this end, a compiler converts the Simulink code to C-code, which can be run natively on the Speedgoat system.

The development of algorithms for hearing devices, such as hearing aids or hearing-aid glasses with a focus on feedback cancellation and beamforming, is one research focus of Prof. Puder's Audio Signal Processing Group. For beamforming, the two microphones in each of the left- and right-worn hearing aids are combined to a four-microphone beamformer system; whereas, in hearing-aid glasses, several microphones can be integrated in the ear pieces. Here, even narrower beams can be realized due to larger microphone distances and a higher number of microphones.

Such hearing systems need to be evaluated under realistic conditions, i.e., worn on the head and not in a free field. We use the Knowles Electronics Manikin for Acoustic Research (KEMAR) [11] as a well-established head model. The KEMAR is shown in Figure 9. This allows us to model the head with respect to head shading as well as to model the ear channel, which is necessary for realistic feedback tests of hearing-aid devices.

FIGURE 8. The real-time Speedgoat system (left) and KEMAR (right) at the Audio Processing Lab.

Several loudspeakers positioned at different AOA serve as sound sources. The real-time system is built in a typical office room, showing a rather high reverberation time. Acoustic curtains, which can be opened and closed, allow tests of the systems within different reverberant acoustic environments.

The Bioinspired Communication Systems Lab

The Bioinspired Communication Systems Lab of Prof. Heinz Koeppl conducts research in statistical signal processing and machine learning in the context of biomolecular systems. Due to the availability of wet-lab facilities in the group (see Figure 10), students can generate their own data by performing single-cell experiments involving fluorescence microscopy and microfluidics, which is shown in Figure 11. The microfluidic chips used for student projects are further optimized and fabricated in-house. The hands-on work also involves the processing of raw imaging data to obtain accurate segmentation and temporal tracking of single cells.

The Receive Beamforming Lab

Prof. Marius Pesavento's Communication Systems Group offers a student experiment in the field of multiantenna receive beamforming. The experiment is based on the WARP v3 Kit by Mango Communications [12] (see Figure 12), which includes an easy access MATLAB interface. The main idea of the experiment is to give students insight into the application of receive beamforming as part of a complete transmitter-receiver chain, starting from the antennas that were designed at TU Darmstadt specifically for the experiment and ending with digital baseband signal processing algorithms implemented in MATLAB. A main challenge in the design of the experiment was to find the best tradeoff between performance and complexity on the one hand, and comprehensibility of the exercise on the other.

In the student experiment, an antenna array is used that consists of eight patch antennas designed for the 2.4-GHz industrial, scientific, and medical (ISM) band on the receiver side (see Figure 13). The antenna array is connected to two WARP v3 boards with four RF ports per board. On the transmitter side,

FIGURE 9. A hearing-aid dummy worn on the ear model of the KEMAR. The dummy is connected via cables to the real-time system.

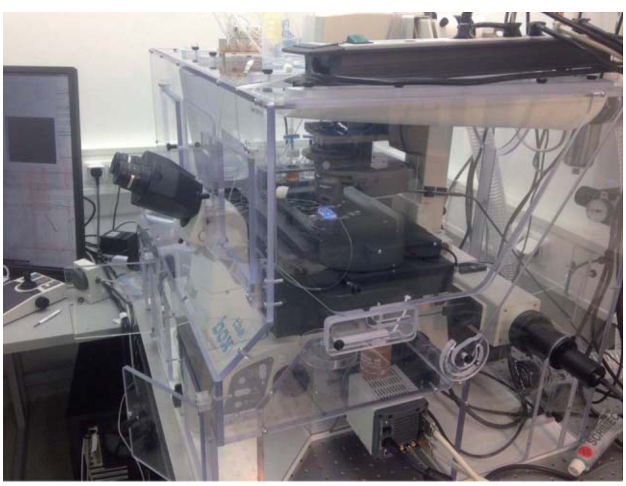


FIGURE 10. The wet-lab facilities at the Bioinspired Communication Systems Lab.

FIGURE 11. Single-cell recording and tracking in a swarming assay of bacteria *Bacillus subtilis* taken at the Bioinspired Communication Systems Lab.

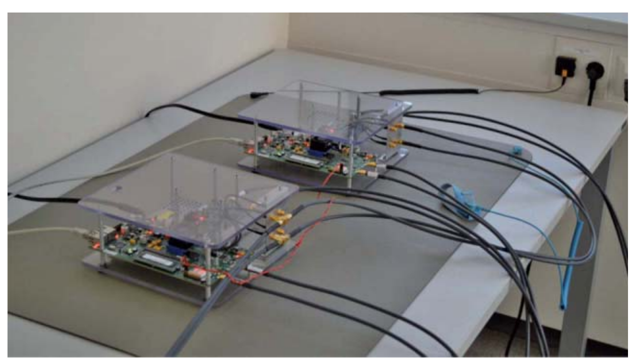


FIGURE 12. The WARP v3 test bed by Mango Communications.

two independent and freely movable patch antennas, also designed for the 2.4-GHz ISM band, are employed. Both patch antennas are connected to a third WARP v3 board and can be operated independently to model two independent transmitters. For simplicity, all boards are synchronized in RF and sampling frequency by external cables and the boards are connected to a MATLAB server via Ethernet, which is used for offline baseband processing.

The goal of the experiment is to provide user data separation by means of receive beamforming. Two different concepts are implemented and tested by the students. In the first approach, receive beamformers are designed based on channel state information acquired from pilots; while, in the second approach, a line-of-sight model is employed to model the channels parametrically. For ease of implementation, a simple MATLAB interface is provided that students can use to perform all the required signal processing, i.e., pulse-shaping, sampling, timing synchronization, and channel estimation. As a result, students can directly focus on the beamforming implementation.

A particular challenge in carrying out student experiments on the wireless test bed is to visualize and evaluate the effects of different procedures and methods. Therefore, the experiment is divided into different tasks. For example, displaying the quadrature phase-shift keying (QPSK) signal constellations before and after demodulation or displaying a spatial power spectrum to estimate the user locations. During the experiment, students can directly see, e.g., the effects of the antenna orientation on the quality of the demodulated QPSK or the variation of the spatial spectrum as the user locations are changed.

In summary, during the course and in the course evaluation, very good feedback was received from the students. The implementation of a complete transmitter-receiver chain helps students to better understand and link the individual operations in wireless communications while the experiments help to visualize the effects of different operations.

Competitions

The SPG seeks to participate in student competitions, as we believe this is one of the best ways to provide undergraduate students with the opportunity to have hands-on experience and to put their signal processing knowledge into practice in a real-world project. Furthermore, students again learn to cooperate within a team and develop their interest in signal processing research. From our experience, the students who took part in competitions show not only higher technical understanding but also higher motivation and enthusiasm. They are inspired by their hands-on experience and their voluntary and ungraded achievements, which can also lead to better overall performance in their studies.

One important aspect to mention again is teamwork. If the team works harmoniously and everyone enjoys their time, creativity skyrockets. When there is a successful competition outcome, students gain additional benefits by having the possibility of visiting a conference, receiving a prize, or gaining prestige. Next, two examples of successful participation in students' competitions are given.

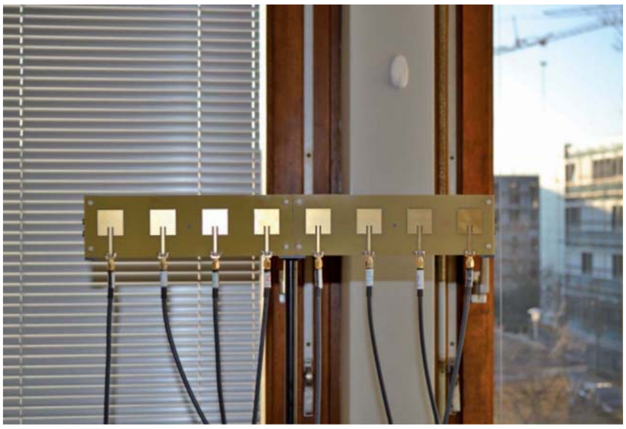


FIGURE 13. The antenna array consisting of eight patch antennas designed for the 2.4-GHz ISM band.

From our experience, students who took part in competitions show not only higher technical understanding but also higher motivation and enthusiasm.

Case study competition by Rohde & Schwarz

Together with the German Association for Electrical, Electronic, & Information Technologies (VDE), Rohde & Schwarz organizes an international case study competition for undergraduate students [13] in the field of mobile communications. Its aim is to offer students the opportunity to expand their scientific knowledge and solve real-life technical problems. It is the organizer's intention that students not only deploy specific theoretical knowledge but also enhance their teamwork and creativity skills. The first round of the competition takes place at universities, where participants work on a technical problem in the area of mobile communications. A jury consisting of one professor and several company employees decides on the best solution. The winning team members are then invited to the finals at the company's headquarters in Munich. In the final competition, teams from different universities compete against each other in finding the best solution to a complex problem. The winning team receives a prize as well as €2,000 for their university.

In 2012, the student team, Shannon's Hounds, of the SPG took part in the case study competition whose theme was "Engineer the future! The future of mobile communications is on you" (see Figure 14). In the competition in which 220 students from Germany and Singapore participated, students had to solve complex tasks concerning the ISO/OSI-layers of the LTE cellular network. At the finals in Munich, modern measurement equipment provided by Rohde & Schwarz had to be used to find solutions. "We were excited to work with real modern measurement devices and have hands-on experience at the finals in Munich," said Mark Ryan Balthasar from

FIGURE 14. Thomas Rösner (far left) and Stefan Diebenbusch (far right) from Rohde & Schwarz stand with team Shannon's Hounds members: (from left) Lisa Hesse, Patricia Binder, Mark Balthasar, Fabio Nikolay, and Cevin Sehrt at the ceremony in Munich, 2012. (Photo courtesy of Rohde & Schwarz.)

the team representing TU Darmstadt. Shannons's Hounds performed outstandingly well and won the competition against ten teams from Germany and Singapore. Each team member received an Apple iPad and EUR€2,000 for the university. The team decided that the prize money would be spent on hands-on experiments for teaching purposes. Two out of the five members of the team are now RAs with the SPG, and one is with the Communications Engineering Group.

The IEEE Signal Processing Cup

The IEEE Signal Processing Cup was initiated in 2013 by the IEEE Signal Processing Society to increase students' interest in signal processing and to raise their awareness of its applications in real life [14]. Undergraduate students are provided with an opportunity to form teams and work together to solve a challenging and interesting real-world problem using signal processing techniques and methods. In 2014, approximately 100 undergraduate students from all over the world took part in 25 different teams. The theme for the first competition was "Image Restoration/Super-Resolution for Single Particle Analysis."

After the release of the new competition theme for 2015, the SPG decided to take part in the second edition of this prestigious competition with its student team, Signal Processing Crew Darmstadt. The task of the competition was to estimate the heart rate using PPG signals recorded from subjects' wrists during physical exercise. See [15] for more information on the competition. RAs Michael Muma and Tim Schäck recruited seven students with interest and motivation in signal processing in August 2014. In total, approximately 270 undergraduate students split among 66 teams registered for the competition.

"For the next months, we arranged regular meetings where we discussed and developed different approaches," remembers Tim Schäck. "We built a biomedical laboratory with our own PPG sensors to be able to take measurements and collect additional data. For this, we employed a student with the necessary hardware skills as an undergraduate research assistant whose task was to develop a framework for the collection of measurements. Hence, the team members also had the option to gain hands-on experience in our lab, which was much help to the students in the competition.

"After submitting our algorithm and results in February 2015, we were more than happy to find out one month later that our team was among the best three teams and that we were invited to take part in the final competition at ICASSP 2015 in Brisbane, Australia. Fortunately, we managed to get five students to fly to Brisbane and present the work at the finals. As if this was not enough excitement for the students, their presentation convinced the jury of the new method and Signal Processing Crew Darmstadt won the IEEE Signal Processing

Cup 2015 against tough competitors from Bangladesh University of Engineering and Technology and Soongsil University in South Korea," says Schäck. Figure 15 shows the members of the team at the ceremony in Brisbane. The competition topic and results have been published in [14].

"After a slow start, we managed to sit together as a team and were able to formulate subproblems, which were assigned among the team members. We discussed several approaches and often ideas that seemed right in the beginning were dismissed or modified to produce even better results," reports team member Maximilian Huettenrauch. "I think the SP Cup can be very helpful in the sense that one can work on close-to-real-world problems. The problem was not as contrived as university tasks often are, and the data was collected from real experiments. It also showed that, often, it is not the most complex and sophisticated concepts that lead to good results, but rather starting out with a basic idea and adding bits and pieces to this initial idea," Huettenrauch continues. Additional feedback from participating students and supervisors have been published in [15]. The main part of our prize-winning algorithm [16] was published at EUSIPCO 2015 in Nice, France, by the two supervisors and one of the undergraduate students, who also continued working on heart rate estimation in his master's thesis project.

Practical remarks for successful design projects

We conclude this article by briefly sharing our experience on some fundamental aspects that we have found to be important for the success of signal processing design projects.

Interculturality

Special emphasis should be given to integrating students from other cultures. In [17], Prof. Zoubir describes challenges in having intercultural groups in research. For example, independence in research has a high priority at TU Darmstadt, but some researchers are not used to such kinds of freedom. Thus, misunderstandings can occur among the team members. Similarly, in engineering design projects, students often work in intercultural teams in which not everyone shares the same work attitudes or values. Here, honest and direct communication is very important.

Integrating students into social activities of the research group helps to recruit good students.

Gender equality

Gender equality is always a central topic for the success of our design projects. Thankfully, we are supported by gender equality representatives who act on behalf of all female students and staff members in the department in all matters relating to research or teaching and also provide other services. Currently, at the SPG, more than half of our RAs are female, which helps attract female undergraduate students to signal processing hands-on projects.

Social activities

Integrating students into social activities of the research group helps to recruit good students. These activities may include events such as an end-of-year party and visits to collaboration partners in research or industry. During such visits, undergraduate students see signal processing in action and often the participants ask for hands-on topics and wish to perform research projects within our group.

FIGURE 15. Members of Signal Processing Crew Darmstadt at the final presentation at ICASSP 2015 in Brisbane, Australia.

Mentorship

From the beginning of their academic studies, our students are mentored by the professors who provide guidance throughout the entire duration of study and offer one-on-one meetings. These meetings also offer the possibility to inform the interested students about the hands-on projects that we offer.

Evaluation

We constantly try to improve our courses, labs, and seminars. For this to happen, we run evaluations by means of detailed questionnaires. Some of the questions explicitly concern hands-on experiences. In this way, we receive and are able to take into account feedback from the students on how to increase the number and quality of hands-on projects.

Acknowledgments

We would like to thank Christian Steffens, Prof. Henning Puder, Prof. Heinz Koeppl, Prof. Marius Pesavento, Ann-Kathrin Seifert, Dr. Michael Fauß, and Mark Ryan Balthasar from the Institute of Telecommunications for their efforts in helping us give this article a broadband overview of the hands-on activities in signal processing at TU Darmstadt. The work of Dr. Muma was supported by the project HANDiCAMS, which acknowledges the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number 323944.

Authors

Tim Schäck (schaeck@spg.tu-darmstadt.de) received the Dipl.-Ing. degree in electrical engineering and information technology from Technische Universität Darmstadt, Germany, in 2013, where he is currently working toward his Ph.D. degree in the Signal Processing Group, Institute of Telecommunications. His research focuses on biomedical signal processing. He was the supervisor of the Technische Universität Darmstadt student team that won the international IEEE Signal Processing Cup 2015 with the competition topic "Heart Rate Monitoring During Physical Exercise Using Wrist-Type Photoplethysmographic Signals."

Michael Muma (muma@spg.tu-darmstadt.de) received the Dipl.-Ing. degree in 2009 and the Ph.D. (Dr.-Ing.) degree (summa cum laude) in electrical engineering and information technology in 2014, both from Technische Universität Darmstadt, Germany, where he is currently a postdoctoral fellow in the Signal Processing Group. His research is on robust statistics for signal processing; biomedical signal processing; and robust distributed detection, classification, and estimation for wireless sensor networks. He was the supervisor of the Technische Universität Darmstadt student team that won the international IEEE Signal Processing Cup 2015. He coorganized the 2016 Joint IEEE Signal Processing Society and EURASIP Summer School on Robust Signal Processing.

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. He then joined the Curtin University of Technology, Australia, as a professor of telecommunications and was interim head of the School of Electrical and Computer Engineering from 2001 to 2003. Since 2003, he has been a professor of signal processing at Technische Universität Darmstadt, Germany. He is an IEEE Distinguished Lecturer (class of 2010–2011), past chair of the Signal Processing Theory and Methods Technical Committee of the IEEE Signal Processing Society, and the editor-in-chief of *IEEE Signal Processing Magazine* (2012–2014). His research interest lies in statistical methods for signal processing applied to telecommunications, radar, sonar, car engine monitoring, and biomedicine. He has published more than 400 journal and conference papers in these areas. He is a Fellow of the IEEE.

References

[1] G. C. Orsak and D. M. Etter, "Collaborative SP education using the internet and MATLAB," *IEEE Signal Process. Mag.*, vol. 12, no. 6, pp. 23–32, 1995.

[2] D. M. Etter, G. C. Orsak, and D. H. Johnson, "Distance teaming experiments in undergraduate DSP," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, 1996, vol. 2, pp. 1109–1112.

[3] J. H. McClellan, R. W. Shafer, and M. A. Yoder, *DSP First: A Multimedia Approach*. Englewood Cliffs, NJ: Prentice Hall, 1997.

[4] J. E. Greenberg, B. Delgutte, and M. L. Gray, "Hands-on learning in biomedical signal processing," *IEEE Eng. Med. Biol. Mag.*, vol. 22, no. 4, pp. 71–79, 2003.

[5] H. T. Wu, P. C. Hsu, C. Y. Lee, H. J. Wang, and C. K. Sun, "The impact of supplementary hands-on practice on learning in introductory computer science course for freshmen," *Comput. Educ.*, vol. 70, pp. 1–8, Jan. 2014.

[6] J. Dagdagan, M. Muma, and A. M. Zoubir, "Robust testing for stationarity in the presence of outliers," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, 2014, pp. 3464–3468.

[7] F. Strasser, M. Muma, and A. Zoubir, "Motion artifact removal in ECG signals using multi\resolution thresholding," in *Proc. 20th European Signal Processing Conf.*, 2012, pp. 899–903.

[8] M. Muma, D. R. Iskander, and M. J. Collins, "The role of cardiopulmonary signals in the dynamics of the eye's wavefront aberrations," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 2, pp. 373–383, Feb 2010.

[9] D. R. Iskander. (2016, Mar. 9). Teaching philosophy [Online]. Available: http://dri.pwr.edu.pl/teaching-philosophy

[10] Speedgoat GmbH. (2016, Mar. 9). Speedgoat: x86/FPGA real-time hardware for Simulink [Online]. Available: https://www.speedgoat.ch/

[11] G.R.A.S. Sound & Vibration A/S. (2016, Mar. 9). KEMAR: Rejuvenated. [Online]. Available: http://kemar.us

[12] I. Mango Communications. (2016, Mar. 9). WARP v3 Kit. [Online]. Available: http://mangocomm.com/products/kits/warp-v3-kit

[13] Rohde & Schwarz. (2016, Mar. 10). R&S Engineering Competition 2015. [Online]. Available: http://karriere.rohde-schwarz.de/en/career/fallstudienwettbewerb/start/

[14] K. M. Lam, C. O. S. Sorzano, Z. Zhang, and P. Campisi, "Undergraduate students compete in the IEEE Signal Processing Cup: Part 1," *IEEE Signal Process. Mag.*, vol. 32, no. 4, pp. 123–125, July 2015.

[15] Z. Zhang, "Undergraduate students compete in the IEEE Signal Processing Cup: Part 3," *IEEE Signal Process. Mag.*, vol. 32, no. 6, pp. 113–116, Nov. 2015.

[16] T. Schäck, C. Sledz, M. Muma, and A. M. Zoubir, "A new method for heart rate monitoring during physical exercise using photoplethysmographic signals," in *Proc. 23th European Signal Process Conf.*, Aug. 2015, pp. 2666–2670.

[17] E. Keller, "Geduld und Moderationskompetenz–Fingerspitzengefühl und traditionelle Kniffe wirken in multinationalen Forschungsgruppen," *Hoch3. Die Zeitung der TU Darmstadt*, vol. 2015, no. 1, Feb. 2015.

SP

Qing Zhuo, Yanpin Ren, Yongheng Jiang,
and Changshui Zhang

Hands-On Learning Through Racing

Signal processing and engineering education through the China National Collegiate Intelligent Model Car Competition

MAIN IMAGE ©ISTOCKPHOTO.COM/SKYNESHER; SCREEN IMAGE ©ISTOCKPHOTO.COM/ALENGO

The Intelligent Model Car Competition (IMCC) of China is an annual collegiate contest where student teams design, build, and race a model car around a track, and the fastest car that completes the track without failure wins [1]. The IMCC is in collaboration with the global NXP Cup Challenge, which was formerly known as the Freescale Cup Challenge until the acquisition of Freescale Semiconductor Inc. by NXP Semiconductors [2]. Creating this smart, autonomous car requires students to develop the hardware and software of motor control to propel and steer their model cars. It provides a collaborative, competitive, and hands-on way for students to learn about and make a synergistic use of theories and techniques from undergraduate engineering studies, such as sensing and control, circuit design and implementation, and embedded system and software programming. The first competition, formerly known as the Smart Car Race, began in 2003 in South Korea with 80 student teams. Since then, the NXP Cup has expanded to China, India, Malaysia, Latin America, North America, and Europe, engaging hundreds of schools and tens of thousands of students a year [2], [3].

China started its nationwide college-level smart car race in 2006. It has undergone a rapid growth since then and has just celebrated its tenth anniversary. The challenges and ingenuity posed by this competition has attracted an increasing number of students year by year. As shown in Figure 1, participation has risen from about 112 teams of 57 colleges in 2006 to over 2,000 teams of more than 400 colleges in recent years. For the past five years, more than 30,000 students have attended the contest annually; and so far this decade-long race has engaged more than 150,000 students in total, providing them a valuable hands-on educational experience of engineering.

As members of the organizing committee of the IMCC, in this article we provide an overview and highlights of the competition tasks and rules, the role that signal processing plays, and the curricula that is built on the competition.

Digital Object Identifier 10.1109/MSP.2016.2619985
Date of publication: 11 January 2017

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

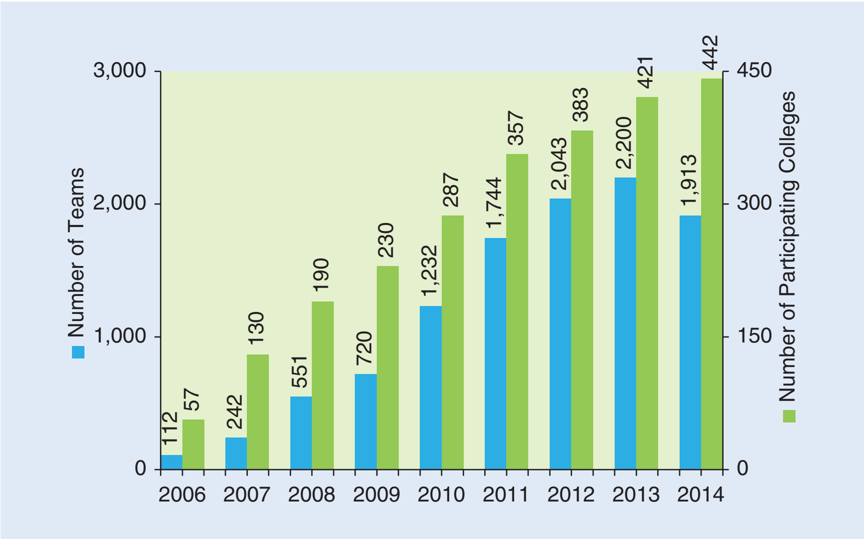


FIGURE 1. The number of participating teams and colleges of the IMCC from 2006 to 2014.

article. It has attracted an overwhelming number of students over the past decade, and its scale is now the largest in the world. What initially began as one competition category has now expanded to six, and the competition tasks have been diversified. Each category has challenges that are suitable for students at a different stage in their college study, so that students ranging from freshmen to seniors can all participate. Along with the IMCC, a large number of microcontroller unit (MCU) teaching labs, textbooks, and innovation training centers have been developed in many universities. The development of such educational material and infrastructure have enabled and expanded hands-on training for engineering students nationwide.

Motivation of a nationwide engineering competition

The launch of the IMCC was supported by the Ministry of Education of China and its Committee of Education Instruction of Automation Specialty. A main motivation to launching the competition was that the engineering curricula at the college level were too theoretical and generally too slow to catch up to the fast pace of the contemporary technological development. As a result, students tended to focus more on test-oriented skills, and did not pay enough attention nor had enough hands-on opportunities to solve real-world engineering problems in a team setting; they would lack curiosity and interest and were not sufficiently motivated to learn and innovate. These problems are not unique in China, as the higher education in engineering in many other countries around the world have faced similar challenges.

China started its nationwide college-level smart car race in 2006. It has undergone a rapid growth since then and has just celebrated its tenth anniversary.

The NXP-sponsored model car competition helps address this problem and bring hands-on engineering education to many college campuses around the world. The IMCC in China has several notable characteristics, including the competition setup, the rules, and the evaluation criteria that will be discussed later in this

Tasks and rules of the model car competition

All racing teams use a standard kit of model car designated by the organizing committee. Team members are required to design and develop their own hardware and software for their cars [4], [5]. As mentioned previously, each finished model car must be capable of self-navigating along a challenging racetrack as fast as it can. The teams will be ranked according to the time taken by the model car to complete one round of the racetrack.

Only undergraduate students are permitted to participate in this nationwide competition in China. Each team is allowed to have up to three students and no more than two faculty advisors. Typically, as shown in Figure 2, an annual competition lasts ten months as an extracurricular activity, from launching in the previous November to the division competition in July, and to the final race in August.

Early rounds of the competition are carried out in eight racing divisions covering different geographic areas in China. The top teams from each race division are qualified for the final race. During the final race, speed-based race sessions are held in which the time for each finalist car to complete one

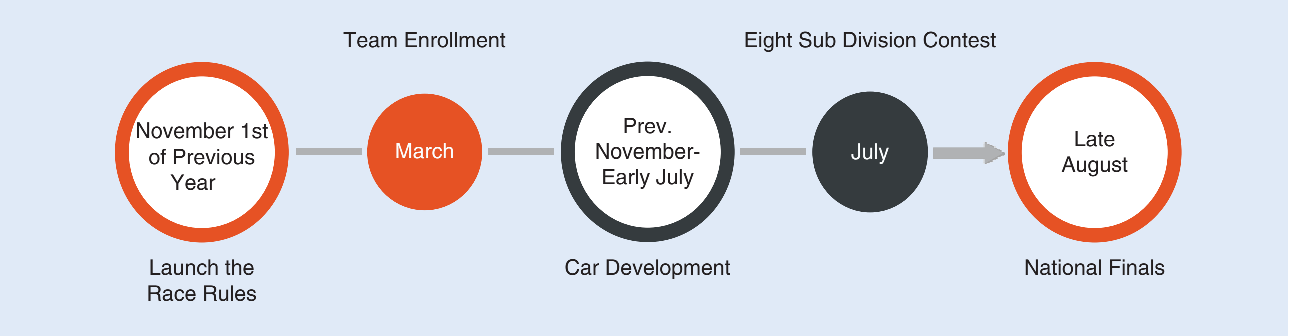


FIGURE 2. The time line of the annual IMCC.

round of the racetrack is used to rank the teams; in addition, an open-ended competition is held concurrently, with innovation themes related to future industrial intelligent cars to encourage students to develop creative ideas and implementations.

Basic elements of the racetrack

With the exception of a beacon-based sensing category to be discussed later in this section, the racetracks are composed of several kinds of elements: straight sections, curved sections, crossroads, hills, and roadblocks (see Figure 3). The characteristics of the racetrack and its elements are given in the rules released at the launch of the competition. A detailed graph of the racetrack is revealed to the teams onsite right before the start of the competition. The sensing and control algorithms embedded in the race cars are expected to work with all these elements and different combinations of them.

What initially began as one competition category has now expanded to six, and the competition tasks have been diversified.

Competition categories

To enable self-navigation of model cars, different kinds of sensors are explored to capture position signals for further processing. Based on the sensor types and race tasks, the competition is divided into several categories that have different levels of technical challenge, as illustrated in Figure 4. The basic categories only require some elementary knowledge of signal processing, control, and circuits, thus allowing younger undergraduate students to participate; on the other hand, the advanced and creative categories may use the technical knowledge and skills from students' design training or capstone projects. The wide variety of categories gives students an opportunity to participate in several competitions during their college career, as they grow in knowledge, experience, and maturity. In what follows, we briefly review the characteristics of each competition category.

Photoelectric sensing

In this competition category, a model car can be equipped with photoelectric sensors, such as an infrared (IR) light-emitting diode sensor and linear charge-coupled device (CCD) sensors, to detect the racetrack. Figure 5 shows a model car equipped with photoelectric sensors. The signals acquired by such sensors are typically binary or one-dimensional. Thus, the algorithms for signal processing and control decisions can be relatively simple compared to the other categories. This category is suitable for younger students with primary engineering knowledge.

Camera sensing

Planar-array CCD or complementary metal–oxide–semiconductor (CMOS) cameras are used to pilot the car in this category. A model car equipped with a camera is shown in Figure 6. Two-dimensional image acquisition and processing requires more computing capabilities than for the other sensing categories. Student teams in this category often equip their model cars with a high-performance 32-bit MCU with larger random access memory (RAM) storage and higher million instructions/second in computing power.

Using image processing and computer vision algorithms, it is possible for a model car to deal with a more complex track layout, predict the direction of the road, and plan for its motion on the racetrack. As a camera can capture images farther ahead of the racetrack, camera cars are usually the fastest among all of the cars in the competition categories.

FIGURE 3. The different elements in a racetrack.

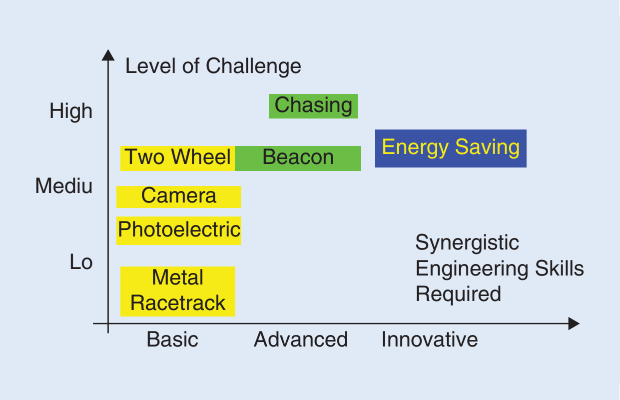


FIGURE 4. The different categories in the competition.

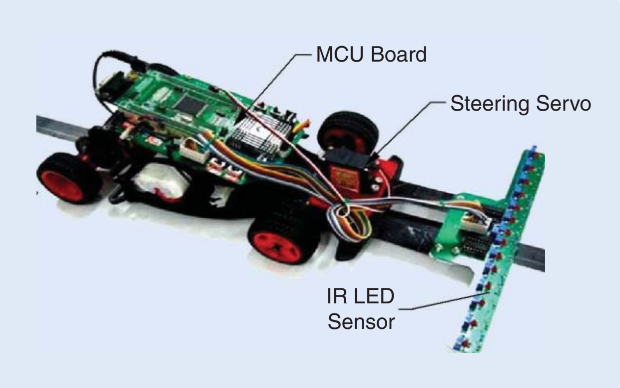


FIGURE 5. A model car with photoelectric sensors.

SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

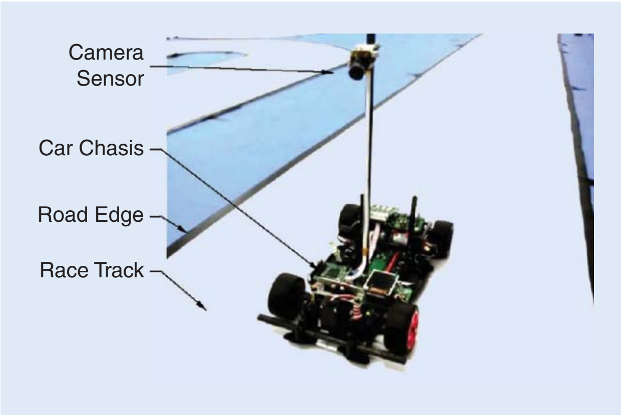


FIGURE 6. A model car with a CMOS camera.

Two-wheel car

There are two specific characteristics and challenges in this competition category. First, the model car only has a total of two wheels, left and right. To propel the car to move upright, sensors such as the gyroscope and accelerometer need to be employed, and the signals from these sensors must be acquired and processed properly in real time. Second, instead of putting visible black edge lines on the racetrack as in the other competition categories, one enameled wire with a diameter of 0.5 mm is laid along the center of the racetrack to guide the movement of the model car. Alternating current (ac) (100 mA, 20 kHz) flows along the wire, which generates an oscillating magnetic field on the racetrack. With such a racetrack design, one way to guide the model car is to use two inductor coils to sense the varying magnetic field.

To sense the racetrack, detection coils can be mounted to the car on a well-designed extending bracket. The battery can also be located low to deliberately lower the center of gravity of the car. As shown in Figure 7, these arrangements can increase the stability and the racing speed of the car.

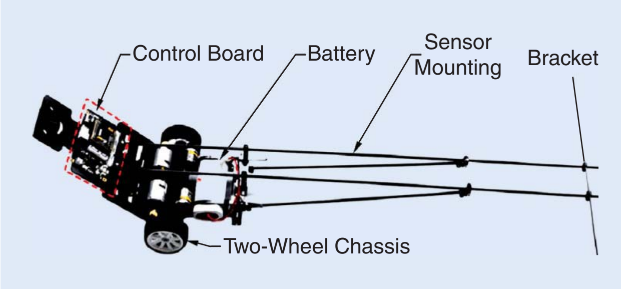


FIGURE 7. A two-wheel car with a long sensor bracket.

Compared to the other competition categories, the signal processing and control methods applied in this category are generally more sophisticated for the race car to maintain its balance while moving smoothly on the racetrack. The Kalman filter algorithm is often used by participating teams to calculate the angle of the car position, and a double closed-loop speed control is implemented to drive the wheels.

Metal racetrack

For the metal racetrack competition category, the racetrack is laid with two parallel strips of aluminum foil, and a direct current (dc) power of 12 V 5 A is applied to the two aluminum strips. A model car participating in this competition category is required to guide itself by detecting the metal foils, and is also allowed to pick up electricity through the aluminum foils to drive its motor. The racetrack components are shown in Figure 8. The main competition goal of this category is to design a highly power-efficient intelligent model car. The judging criteria is not only by the speed of the car, but also by the total energy consumed during the competition.

To sense the guiding metal strips, several coils may be mounted in front of the race car, and a high-frequency ac signal from an oscillator circuit may be applied on the coils. The alternating electromagnetic field generated by the coils will induce the eddy current on the surface of the metal strips. In turn, the eddy current will change the amplitude and frequency of the ac in the coils. The specific variation depends on the related distance between the detecting coils and the metal strips. By using the amplitude demodulation or frequency demodulation, the car can detect its deviation from the guiding strips. Such sensing and signal processing methods may be implemented by circuit or by software.

Two-car chasing

Sensors used in this competition category are similar to the one for camera sensing. Here, each team is required to design two model cars to run one after another on the racetrack. Figure 9 shows two students preparing their chasing cars on the racetrack.

The final score (T) contains two portions: the total time ($T1$) for the two cars to complete a round, and the lag ($T2$) between them as they cross the finishing line. The formula for calculating the total score is $T = T1 + 5T2$.

To perform well in this competition, the two race cars should coordinate well by wireless communication. They must avoid colliding with each other while at the same time avoid being too far apart. This is the most challenging among all the competition categories.

Detecting the distance between the two cars is crucial in this competition.

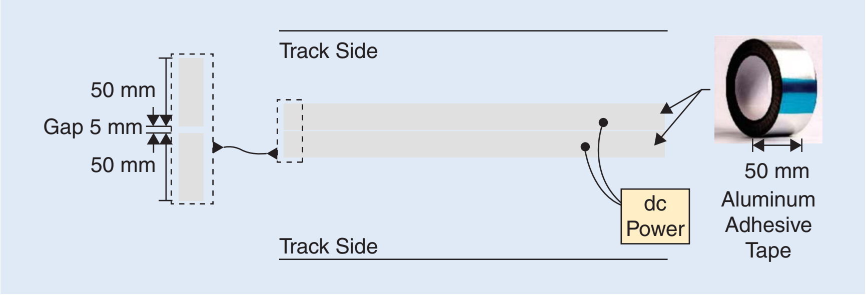


FIGURE 8. A racetrack option with a metal strip.

Usually, an ultrasonic signal and an IR signal are sent back simultaneously from the leading car. The second car can determine the distance by detecting the time-lag between the receiving ultrasonic and IR signals.

Beacon-based racing

Unlike the other competition categories, there is no visible racetrack in this competition. Several beacons are distributed on the competition field and can be turned on by a referee system in a random sequence. Model cars should move to approach the lighting beacons. As long as a car moves inside the detection region of a beacon, the referee system will turn on the light of the next beacon. The arrangement of the beacon group field is illustrated in Figure 10.

The challenges of this competition include detecting targets that may be located relatively farther away from the race car, avoiding collision onto the beacon during the navigation, and planning a motion path for the race car.

Many participating teams have used a camera to search for the beacon. To differentiate from the surrounding light sources, the beacon flashes at about 10 Hz. This flash pattern is used to locate the beacon in the middle of ambient light. Participants have shown that target detection based on the frame difference image is a robust approach to locate the beacon in most types of environments.

Forward-looking innovation category

To foster creative thinking, the IMCC also has an open-ended category inviting teams to contribute innovative ideas and designs. Different themes are set up for each year. The most recent competition's theme, for instance, was innovative designs on energy saving and a future smart city. Teams are required to submit a detailed technical report as well as a video of their work. The top teams are invited to the final race each August to showcase their work. Figure 11 shows a winning entry of a recent competition, where the student team designed and prototyped a parking facility in a futuristic community.

FIGURE 9. Students preparing two chasing cars on the racetrack.

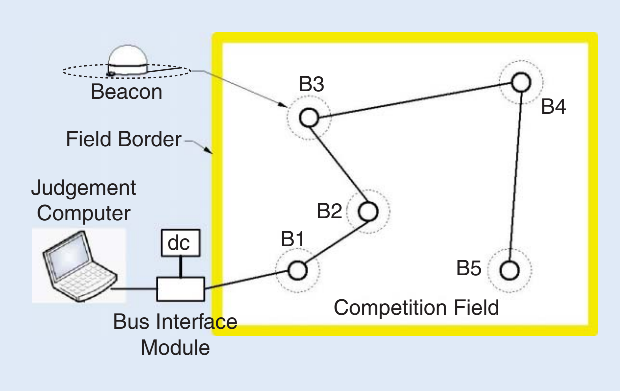


FIGURE 10. The beacon group competition field.

FIGURE 11. An exhibit of the creative design category of the IMCC.

Signal processing techniques used in the competition

There are many different technologies used in a model car to compete in the respective category [6], and signal processing is one of the key components. Students receive hands-on training and strengthen their understanding of signal processing through the competition.

Signal sensing and sampling

As the model car is controlled by the MCU, almost all signals coming from the sensors would be sampled and converted into digital. Some of these signals are already digitized by the sensor module themselves so that they can be passed directly from the sensor to the data transfer port of the MCU. Some other sensing signals are obtained in an analog form, for example, the photoelectric sensor signal and induction coils signals. These signals need to be sampled after conditioning circuits by an analog-to-digital converter (ADC) module on the MCU.

As students learned from fundamental signal processing courses, the sampling frequency need to be properly determined according to the property of the signal. For example, the signal that comes from the induction coils in the two-wheel competition has an alternate voltage with a fixed frequency of 20 kHz. It is a narrow-band signal, and thus with proper processing, it could be sampled at a frequency that is far lower

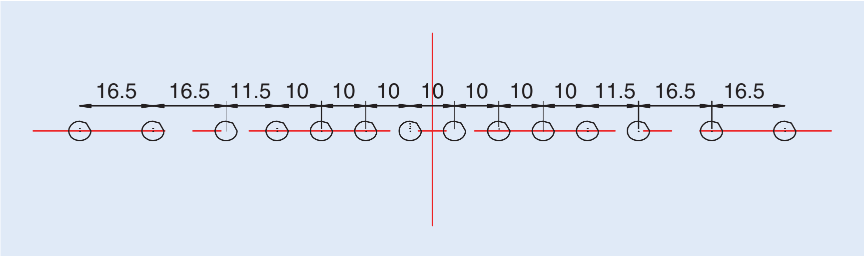


FIGURE 12. The layout of the photoelectric sensor placement.

than the nominal Nyquist frequency. Based on these sampled data, the amplitude of the signals can be determined.

Another example of sampling is to detect the center of the racetrack. In the photoelectric sensing competition, the sensors sample the light intensity of the racetrack surface spatially at several isolate spots. As the number of the sensors is confined to no more than 16 by the competition rules, the arrangement of the sensors must be well designed to increase sampling range and accuracy. Figure 12 shows a common layout of photoelectric sensors in the competition, whereby the sensors are arranged in one line in front of the car with different spacing between them, and the space in the middle is smaller than the outer ones. This kind of nonuniform sampling was shown to perform better than the uniform spatial sampling in the competition.

Denoising and parameter estimation

An important task of the competition is to take the sensing signals that are often noisy or distorted and process them to extract important parameters or other useful information and pass it on to the control and decision mechanism to race the model car.

Student teams have employed a variety of filters in the model car control systems to deal with the interference coming from external sources or onboard circuits. They have used both analog filters to condition the sensing signals, and digital filters implemented by embedded software for linear, nonlinear, or adaptive signal processing.

The main source of interference on board comes from the motor driver circuit with high peak currents up to 20 A. The noise sparks travel through the power supply line into the control circuitry, causing erratic behaviors. Most of the noise can be reduced by using analog filters in the sensing and power circuits. However, some noise is still left in the sensing signal and is sampled into the MCU. This remaining noise can be dealt with by digital signal processing.

Other noise sources may also be seen in each specific competition setup. For example, we previously mentioned that in the two-wheel competition, a model car must keep an upright pose while running along the racetrack. An inertial measurement unit (IMU) is mounted on the chassis to measure the inclination of the model car. The IMU outputs two signals: the gyroscope signal that gives the angular speed, and the accelerometer signal that provides information on the "down" direction. Although the dip angle of the car can be calculated according to the accelerometer signal alone, the movement of the car produces much noise mixed with the angular position. To address this problem, student teams have employed Kalman filters when computing the angular position (shown in Figure 13), combining the gyroscope and accelerometer signals in a denoised fashion. Some student teams have simplified the filtering algorithm to allow for more efficient estimation of the attitude angle of a two-wheel race car. In addition, one of the teams in the two-wheel competition carefully studied the noise levels in the IMU data when the pulse-width modulation (PWM) voltage with different duty ratio is applied to drive the motors. They incorporated this into a Kalman filter implementation to further reduce the noise when estimating the pose of the model car.

Two-car chasing is the most challenging among all the competition categories.

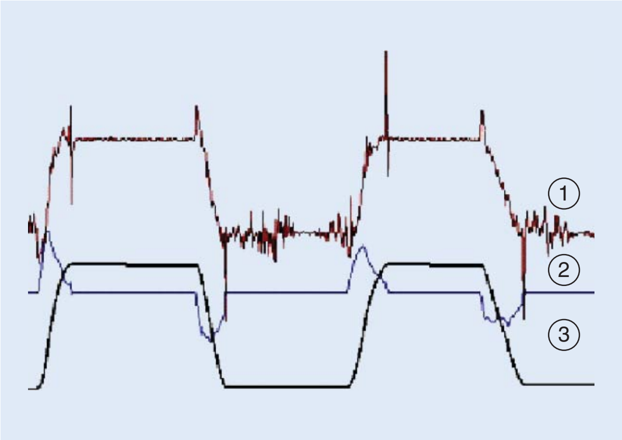


FIGURE 13. Using a Kalman filter to determine the angular position of a car. Curve 1: accelerometer signal; curve 2: gyroscope signal; and curve 3: the Kalman filter output signal.

Image processing and computer vision

Thanks to the power of visual sensing, student teams prefer using cameras to guide their model cars in the competition categories of camera sensing, chasing, and beacon-based racing. By using a camera to continuously capture images of the racetrack, a wide range of information about the car's current state and potential future situation can be inferred.

Detection range is a key factor affecting the achievable speed of a self-navigating race car. Because there exists a constant delay when a servo changes the direction of the front wheels of a race car, detection in advance can compensate this delay. The sooner the detection of the curve on the racetrack is, the higher speed the race car can run without going out of the bounds of the racetrack. Image processing from camera data also enables the extraction of more detailed information of the racetrack. For example, detecting the obstacles on the racetrack, including hills and roadblocks, is generally easier and more accurate by a camera than by other types of sensors.

Despite the aforementioned advantages, there are challenges in performing image processing on a low-cost embedded platform. The main limitation comes from the shortage of RAM memory capacity and the limited computing power of the MCU. To strike a good tradeoff between the processing speed and accuracy, it is desirable for the MCU to subsample the image into a lower resolution to work with, and such low-resolution low-quality image poses a challenge for student teams to reliably detect the racetrack and other elements on the track.

Detection range is a key factor affecting the achievable speed of a self-navigating race car.

Figure 14 shows an example scene of a model car on the racetrack. Inset (a) is the binary image captured in the MCU's RAM. Inset (b) shows the center line of the racetrack extracted from the image above by the image processing algorithm.

The image captured by the camera on board of a race car also has a large distortion caused by the low viewing angle of the camera. To obtain more accurate information of the road, a perspective transformation is applied upon the image to correct such distortion. Figure 15 shows the effect of perspective transformation. Figure 15(a) reveals the original grayscale image; (b) shows the center line of the racetrack extracted from the gray image; and (c) gives the correct center line after the perspective transformation.

In summary, signal processing techniques are applied extensively on various aspects of the race cars. This competition allows students to practice their knowledge of signal processing in task-oriented training scenarios, which complements the traditional classroom learning to better prepare students for future real-world R&D.

Curricula development based on the competition

An important goal of the IMCC is to integrate the learning of multidisciplinary knowledge and the training of comprehensive abilities important to engineering professionals through the process of building and racing model cars. As mentioned earlier, the respective expertise involved include analog/digital electronics, embedded system, electronic design automation, control engineering, signal processing, pattern recognition, and more. Figure 16 illustrates the overall process for students to attend the competition, together with the respective disciplines supporting the different tasks in the competition.

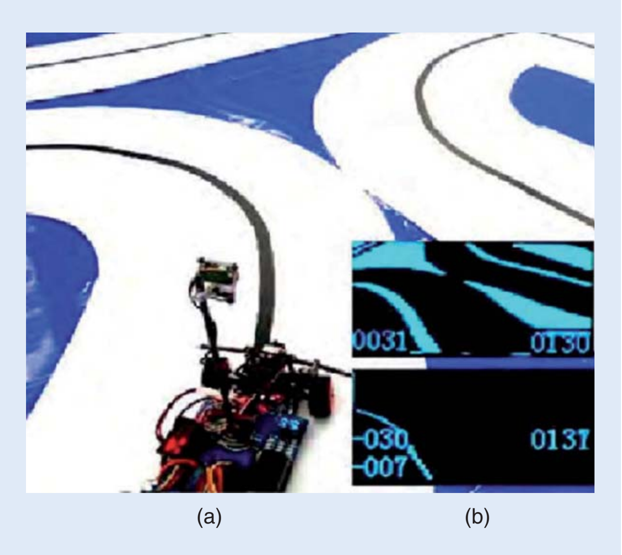


FIGURE 14. A model car with a camera mounted on the top of the pole in the middle of the car. The inset (a) shows the binary image captured in the RAM; (b) shows the center line of the racetrack.

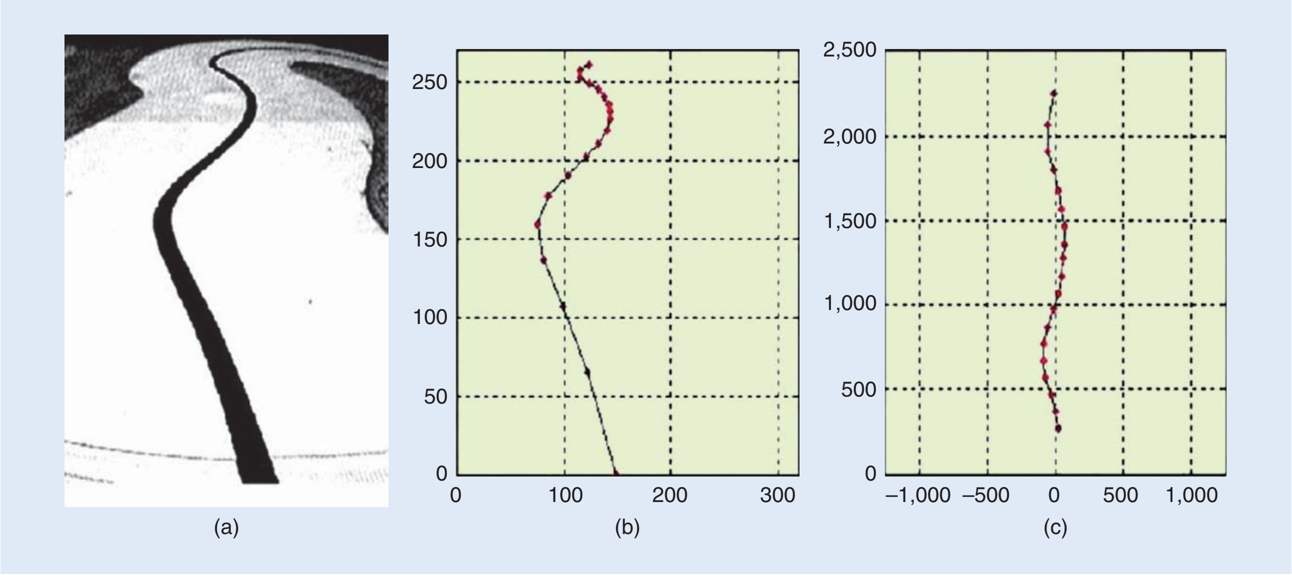


FIGURE 15. (a) The origin image of the S-shape curved racetrack. (b) The center line of the racetrack. (c) The center line after the perspective transformation.

After a decade of annual intelligent car competitions in China, a total of 176 practical training centers have been set up in 132 colleges across the country, and 115 courses have been developed or redesigned in Chinese universities. More than 36 textbooks have been published with different focuses and scopes, providing guidance on the core knowledge and skills to students who are interested in participating in the race car competition, or simply for a fun hands-on extracurricular practice. In addition, a large number of technical reports on the successful approaches in the previous competitions have been archived, from which new students can gain insights on the essential functional blocks and the associated circuitries.

More than 150,000 students have attended the competition and flourished because of it.

Furthermore, an excellent group of faculty members and instructors nationwide have served as mentors to guide student teams. They provide guidance and technical resources to students, while leaving students room to think independently, make design decisions, and implement the designs. When students encounter problems in hardware or software design and testing, mentors can guide them to troubleshoot the problems. They can also advise students to evaluate the possible outcomes of a design decision. With help from these mentors, students can learn how to balance analysis and experimentation, which helps enhance their theoretical understanding and practical execution capabilities. Also, mentors usually discuss and summarize with their students the experiences and lessons learned after each competition.

Closing remarks

The IMCC in China has received many awards and recognitions in recent years. By providing college students and programs with an engaging way to learn/teach, this annual competition has built a strong reputation and fostered

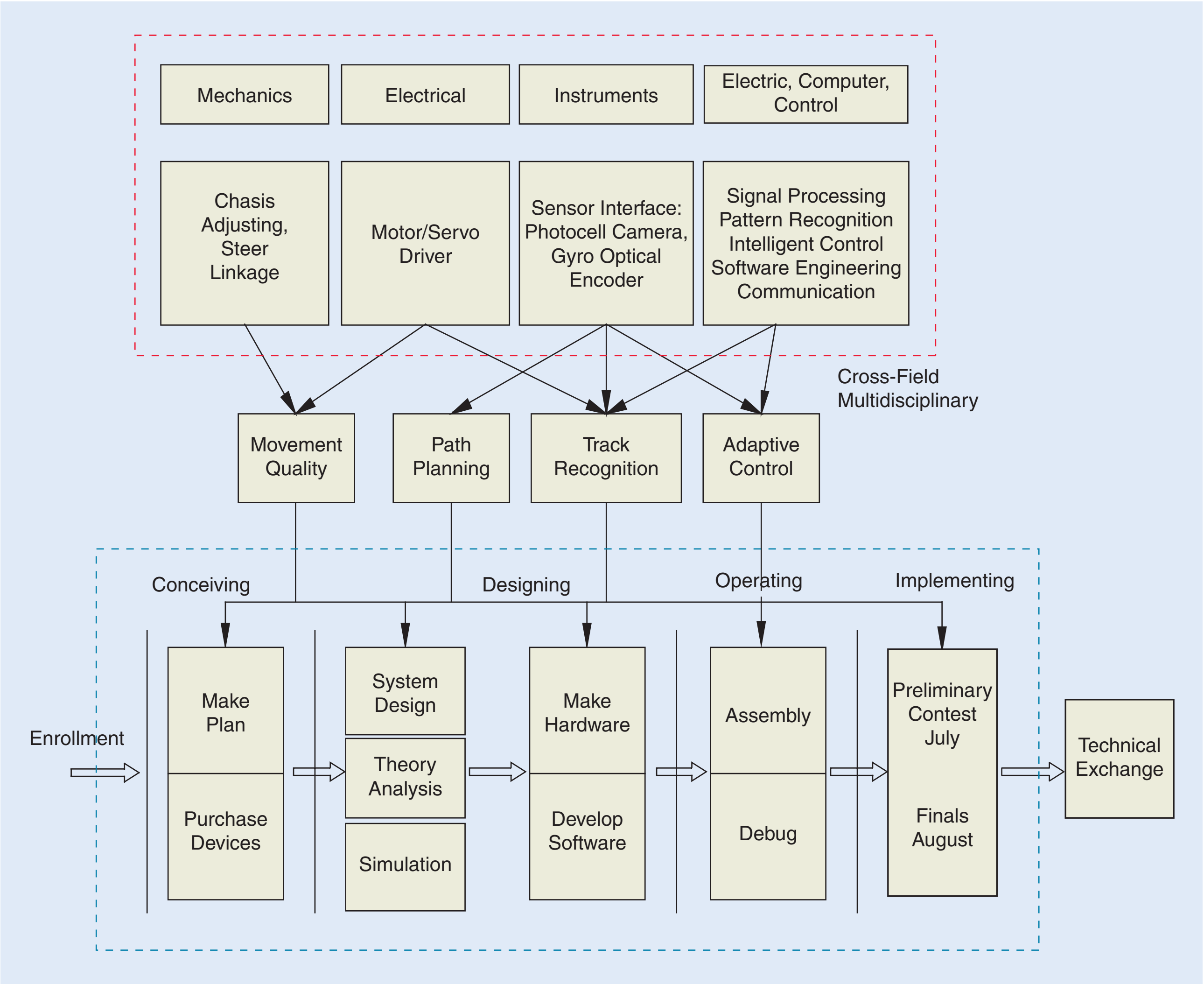


FIGURE 16. An illustration of multidisciplinary involvement and synergy for the IMCC race car design and implementation.

interactions and collaborations with many organizations. The industrial sponsor, Freescale Semiconductor, which merged with NXP, has enthusiastically supported the competition since 2006. More than 150,000 students have attended the competition and flourished because of it. An overwhelming amount of positive feedback from student participants has been received. The IMCC has provided students with an opportunity to learn multiple engineering disciplines, most notably electrical and computer engineering and mechanical engineering, and put them into a synergistic use. The IMCC helps students strengthen communication and teamwork skills and inspires them to pursue engineering careers and become future innovators.

> **The IMCC helps students strengthen communication and teamwork skills and inspires them to pursue engineering careers and become future innovators.**

The competition has successfully met the educational goals originally set by the administrators and sponsors. To ensure its continued success, we are working on addressing several newly emerging issues. For example, several competition categories that use optic sensors to detect the road have seen a nontrivial amount of requirements for the competition conditions for the model cars to work well, and this added an extra burden to the onsite event organization. Also, as the number of teams increases significantly, the cost of organizing the race also increases. To overcome these difficulties, the competition tasks and rules need to be revised periodically. Another new trend with the help of the Internet is the possibility to build some standard competition platforms, which allow students to download their software into their model car remotely, facilitating their participation of the primary competition.

The contents and formats of the IMCC should adapt with new technological advancement. The state-of-the-art technologies and the advanced educational concepts rejuvenate the competition, supporting engineering students to develop their interests and embark on their professional careers.

Acknowledgments

We would like to thank our colleagues on the IMCC committee, including Jingchun Wang, Kaisheng Huang, and Ming Zeng, for their collegial efforts during the organization of the competition and the suggestions in conceiving the competition setup and rules. This work was supported by the National Experimental Teaching Demonstration Center Program, National Special Fund for Improving Basic Education Conditions and NSFC the National Science Foundation of China (NSFC grant 91420203).

Authors

Qing Zhuo (zhuoqing@tsinghua.edu.cn) received his B.S. and Ph.D. degrees in control theory and control engineering from Tsinghua University, Beijing, China, in 1993 and 1997, respectively. He has been serving on the faculty of the Department of Automation at Tsinghua since 1997. His research interests include signal processing and pattern recognition, embedded computer systems, and power electronics. His educational awards and recognitions received from Tsinghua University include Ten Most Favorite Teachers (2010) and Outstanding Mentor (2012). He has been a key member on the Intelligent Model Car Competition Organizing Committee for ten years.

Yanpin Ren (renyp@tsinghua.edu.cn) received her B.S. and Ph.D. degrees in control theory and control engineering from Tsinghua University, Beijing, China, in 1995 and 1999, respectively. She worked at Lenovo from 1999 to 2006 with an R&D focus on smart homes. She is currently a senior engineer in the Department of Automation at Tsinghua University. She is heavily involved in the design and teaching of electronics laboratory courses. Her research interests include digital and analog circuit design, embedded systems, and virtual instruments.

Yongheng Jiang (jiangyh@tsinghua.edu.cn) received his B.S. and Ph.D. degrees in automation from Tsinghua University, Beijing, China, in 1998 and 2003, respectively, where he is currently a faculty member in the Department of Automation. His research interests include process scheduling, trajectory planning, and optimization technology.

Changshui Zhang (zcs@mail.tsinghua.edu.cn) received his B.S. degree in mathematics from Peking University, Beijing, China, in 1986, and his M.S. and Ph.D. degrees in control science and engineering from Tsinghua University, Beijing, in 1989 and 1992, respectively. He joined the faculty of the Department of Automation, Tsinghua University, in 1992, where he is currently a professor and the director of the Institute of Information Processing. He has authored more than 200 papers. His research interests include pattern recognition and machine learning. He is an associate editor of the journal *Pattern Recognition* and served as the associate chair of education in his department and university.

References

[1] Q. Zhuo, J. Wang, K. Huang, Y. Jiang, M. Zeng, G. Shen, and X. Wang, "A ten years review of the national student intelligent car race," in *Collected Works of Decade Development of Collegiate Intelligent Car Racing in China*. China Univ. Teaching, 2012 (in Chinese).

[2] NXP/Freescale Cup model car race overview [Online]. Available: https://community.nxp.com/docs/DOC-1011

[3] The 2016 NXP Cup rules used for competitions in Americas, Malaysia, and Taiwan [Online]. Available: https://community.freescale.com/docs/DOC-1284

[4] G. Z Shen. (2015). Illustration of complex engineering problem: Intelligent mode car competition [Online]. Available: http://www.smartcar.au.tsinghua.edu.cn/column/pxxz

[5] Competition rules of the 11th Intelligent Car Competition in China, 2015 [Online]. Available: http://www.smartcar.au.tsinghua.edu.cn/

[6] J. Sun, W. Huang, and S. Niu. (2009). Technical report documents of the teams participated on the final competition in China [Online]. Available: http://www.smartcar.au.tsinghua.edu.cn/

SP

Erik G. Larsson, Danyo Danev,
Mikael Olofsson, and Simon Sörman

Teaching the Principles of Massive MIMO

Exploring reciprocity-based multiuser MIMO beamforming using acoustic waves

MAIN IMAGE ©ISTOCKPHOTO.COM/SKYNESHER; SCREEN IMAGE ©ISTOCKPHOTO.COM/ALENGO

Massive multiple-input, multiple-output (MIMO) is currently the most compelling wireless physical layer technology and a key component of fifth-generation (5G) systems. The understanding of its core principles has emerged during the last five years, and material is becoming available that is rigorously refined to focus on timeless fundamentals [1], facilitating the instruction of the topic to both master- and doctoral-level students [2]. Meaningful laboratory work that exposes the operational principles of massive MIMO is more difficult to accomplish. At Linköping University, Sweden, this was achieved through a project course, based on the conceive-design-implement-operate (CDIO) concept [3], and through the creation of a specially designed experimental setup using acoustic signals. The course was developed with the following three objectives in mind:

- Exposure of students to emerging concepts and to the technology of the future, not of the past. This target was inherently met via the focus on massive MIMO.
- Promotion of a systems view of thinking, requiring the synthesis of knowledge acquired through the classical curriculum. This goal was achieved through the development of a unique acoustic, reciprocity-based massive MIMO laboratory setup (Figure 1), with students taking the lead both in its development and evaluation and, subsequently, its refinement.
- Fosterage of genuine teamwork, preparing students for a dynamic work environment. This was efficiently facilitated through the adoption of the CDIO project concept, which integrates technical work with the instruction of project management and entrepreneurial skills.

Massive MIMO: The scalable 5G wireless access technology

Massive MIMO exploits the use of large antenna arrays at wireless base stations to simultaneously serve a large number of autonomous terminals through spatial multiplexing. The multiplexing takes the form of beamforming, also known as *multiuser precoding*, effectively creating transmitted signals

Digital Object Identifier 10.1109/MSP.2016.2618914
Date of publication: 11 January 2017

that add up constructively on the spots where the terminals are located and destructively almost everywhere else (Figure 2).

As the ultimately most useful and scalable form of multiuser MIMO, massive MIMO departs from conventional MIMO technology in several ways:

- The acquisition of accurate instantaneous channel state information (CSI) at the base station is facilitated through time-division duplexing operation and the transmission of pilot waveforms by the terminals. On uplink, terminals transmit pilots and payload; subsequently on downlink, the base station beamforms to the terminals. Reciprocity of uplink-downlink propagation is essential to the use of uplink CSI for downlink precoding, and it is achieved in practice through calibration of the radio-frequency (RF) chains [4]. The reliance on reciprocity permits accurate channel training even in highway-speed mobility scenarios.
- Extraordinary spectral efficiencies are achieved [5], although no attempt is made to operate at the Shannon limit. The base station uses linear signal processing, and the terminal uses almost no signal processing at all. CSI is acquired only at the base station, obviating the need for downlink pilots. This renders massive MIMO entirely scalable with respect to the number of base station antennas.
- Each terminal is assigned the full bandwidth. Thanks to channel hardening, the effective channel gain for each terminal is deterministic and frequency- (subcarrier-) independent, which greatly simplifies resource allocation problems, and facilitates simple closed-form solutions for power control.
- Different base stations do not cooperate, other than for "slow" tasks such as power control and pilot sequence assignment. Macrodiversity against shadow fading is accomplished by appropriate terminal-to-base station association; the large number of spatial degrees of freedom guarantees that every base station has room to accommodate additional terminals with very high probability.
- By virtue of the large array gain and the ability to null out interference, max-min fairness power control is feasible and can be exploited to yield uniform quality of service within each cell.

Massive multiple-input, multiple-output (MIMO) is currently the most compelling wireless physical layer technology, and a key component of fifth-generation (5G) systems.

Communications laboratory exercises: RF versus acoustic

Laboratory work that uses communications over RF in general requires substantial equipment investments: the availability of a sufficiently interference-free environment (alternatively spectrum licenses for experimental operation); the use of high sampling rates, which, in turn, generates large quantities of baseband data; and the access to high-performance measurement equipment for calibration and debugging purposes. Owing to precise requirements on the phase reference distribution, these difficulties are significantly accentuated

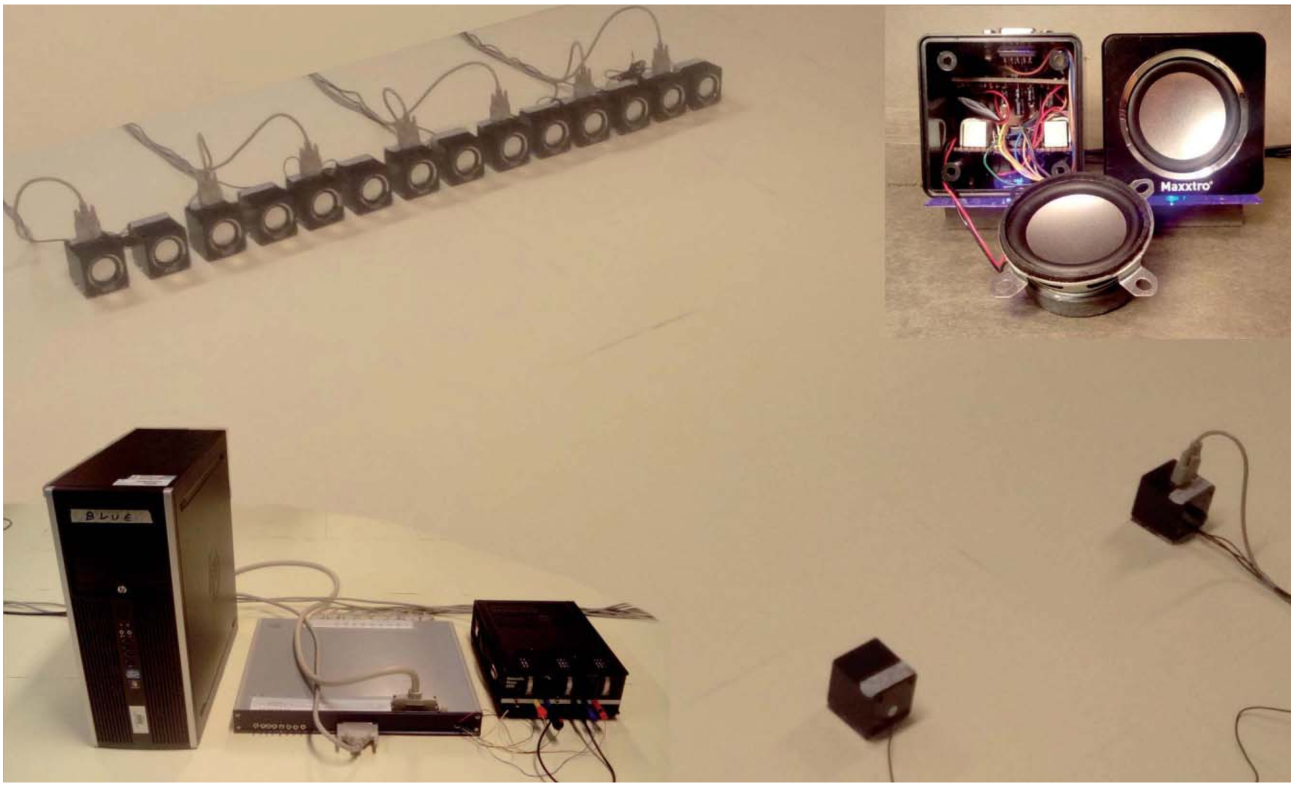


FIGURE 1. The massive MIMO laboratory setup at Linköping University. (Photo courtesy of Mikael Olofsson.)

when going from single-antenna to multiple-antenna setups. Consequently, experimental activities on massive MIMO are concentrated to research-grade test beds that require the investment of many person-years to build, and permanent engineering staff to maintain—thus inherently unsuitable for instructional purposes. Worldwide, only few such test beds, operational and under construction, are known [5]–[10].

The basic physics of wave propagation is substantially the same for electromagnetic wave propagation and for acoustic wave propagation, disregarding polarization aspects. Consequently, wireless communications course laboratory work may use sound waves instead of RF, which for point-to-point communications requires only a loudspeaker as a transmitter and a microphone as a receiver. The wavelength of sound in the audible regime is comparable to the wavelength of radio in the gigahertz regime; hence, the channel coherence time for acoustic communication indoors is sufficiently long to easily permit real-time experimentation over a time-invariant channel. Additionally, many phenomena, such as small-scale fading, can be observed with an acoustic setup as well. The low bandwidth of audible sound results in low sampling rates, modest requirements on clock synchronization, and manageable amounts of baseband data. These features are routinely exploited in many university course laboratory exercises.

Spoofing is accomplished using fake biometric samples expressly synthesized or manipulated to provoke artificially high comparator scores.

Teaching MIMO and massive MIMO principles in the lab

Laboratory work on point-to-point wireless communications is straightforward. In contrast, serious experimentation with massive MIMO concepts is a less obvious task, owing to the need for many simultaneously operating transceivers and the uplink-downlink reciprocity requirement. The experimental setup developed in our course addressed this difficulty as follows:

- A loudspeaker element natively functions as a sound transmitter when fed with an amplified input signal. It was experimentally revealed that oppositely, it also can function well as a receiver (microphone), by connecting it to a high-input-impedance amplifier and measuring the resulting signal.
- A set of off-the-shelf loudspeaker elements was acquired and redesigned as follows:
  1) the built-in amplifier chain was retained, yielding the transmit branch.
  2) an amplifier was installed to use the loudspeaker element as microphone, yielding the receive branch.
  3) a relay that galvanically isolates the transmit and receive branches was installed.

The required additional electronics was fitted inside of the loudspeaker element casing (Figure 1). The rebuilt loudspeaker element, now functioning reciprocally in both the transmit and receive directions, is referred to as a *transceiver device* in the following.

- Fourteen transceiver devices were used to form a base station array and two were used to form two independent terminals. All devices were connected to a personal computer equipped with multichannel digital-to-analog (D/A) and analog-to-digital (A/D) cards. At a 6.25-kHz sampling rate per channel, the aggregate baseband data rate from all channels combined equaled 1,200 kilobits/second, facilitating real-time processing in a MATLAB environment.

For further technical details, see "Detailed Description of the Transceiver Device Design." A live demonstration may be found at [11].

Promoting a systems view of thinking

University education traditionally is rather modularized, consisting of courses with a well-defined but often very specific scope. In contrast, practicing engineers must be able to solve problems by synthesis of knowledge received both in formal training and courses, and acquired from experience. The traditional curriculum in communications is no exception: while much time in class is spent on proving capacity theorems and computing error probabilities, most laboratory time (and most engineering efforts) tends to be spent on "making it work": solving practical problems, and integrating components together into systems. Formal training and knowledge of textbook material is indispensable, but real problem solving requires both trial-and-error, and the sourcing of information from books, colleagues, and the Internet.

The experimental setup and its development trained students in every aspect of communications, ranging from the understanding of wave propagation, fundamental massive MIMO theory, hands-on design and soldering of electronics, software programming, and algorithm design. This substantially took the students' understanding of a "system" from the input-output box typically taught in signals-and-systems courses, to a "system" meaning a large set of connected, diverse components. Answers were not given to the students but had to be sought through consultation with faculty and senior graduate students. In fact, as the project was previously untested, in many cases instructors did not know the answers to all of the questions. A minor risk was taken that the course would not work out, in which case a backup plan was to only perform certain measurements, so that reporting of a minimum level of requirements for the course could be adequately completed.

Providing true teamwork experience through a CDIO project

The CDIO concept [3] was developed by the Massachusetts Institute of Technology in the United States and Chalmers Institute of Technology, Linköping University, and Royal Institute of Technology, all in Sweden, in the late 1990s and the early 2000s. The initial efforts were supported by the

Knut and Alice Wallenberg Foundation. Since then more than 120 schools and universities worldwide have joined the CDIO initiative. One of the reasons for the proliferation and success of the concept is that it profoundly facilitates teamwork training for students.

The CDIO framework focuses on the teaching of engineering fundamentals from the perspective of real-world systems and products. Through constant input from students, faculty, and engineers, the concept has evolved over the years. It currently represents the state of the art in project course organization in engineering schools, preparing the students for the challenges of professional life in industry.

The CDIO framework focuses on the teaching of engineering fundamentals from the perspective of real-world systems and products.

As one of the founders of the CDIO concept, Linköping University has adopted and implemented the idea from its inception. All engineering master's programs include a mandatory CDIO course in the curriculum. In collaboration with several global companies that have local offices in Linköping, notably Ericsson and SAAB, a specific project model, Linköping Interactive Project Steering (LIPS) [12], has been developed to support the project management process. The LIPS model introduces a set of documents that the students are required to write during the project, as well as a set of milestones and tollgates that underpin the planning of the project work. In the beginning of the course, students usually express doubts

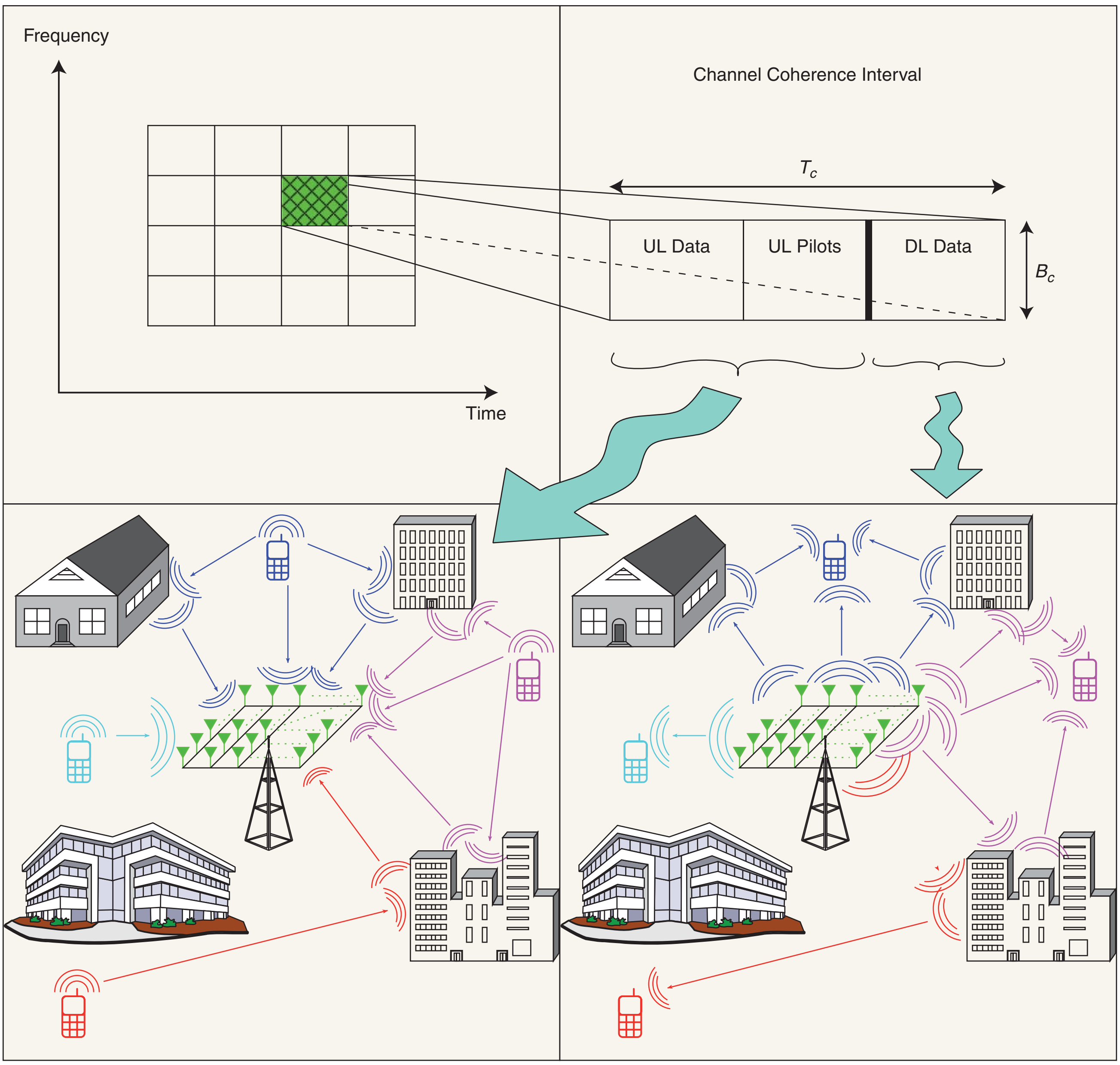


FIGURE 2. Massive MIMO.

Detailed Description of the Transceiver Device Design

The setup consists of a general-purpose computer, digital-to-analog (D/A) and analog-to-digital (A/D) converter cards, modified active loudspeakers, and signal distribution hardware. An overview of this setup is given in Figure S1. Sixteen channels are used from each of the A/D- and D/A-cards (DA12-16 [S1], [S2] and AD12-64 [S3], [S4] from Contec), one A/D- and D/A-channel per speaker. The speakers are used both as speakers and as microphones. We used small active off-the-shelf speakers (Maxxtro MX-US-08 [S5]), where the original power amplifiers are used as such.

The speakers come in pairs, where one serves as the master and contains all the electronics, while the other serves as the slave and contains only a speaker element. The master speaker was modified to facilitate the selection between two modes: either both the master and the slave operate as loudspeakers, or they operate as microphones. This modification is illustrated in Figure S2. On the detection board, relays are used to switch between two operation modes: 1) using the speaker elements of a loudspeaker pair as transmitters, by connecting them to the outputs of the power amplifiers, or 2) using the speaker elements as microphones, by connecting them to the inputs of the receive chain. This way, the hardware supports half-duplex communication.

The receive chain for a single speaker is essentially a traditional instrumentation amplifier that usually is built around three OP-amps and seven resistors to determine the voltage gain, as illustrated in Figure S3. In our implementation (Figure S2), we use five OP-amps. The first two OP-amps were placed on the circuit board in the master speaker and form a differential-in and differential-out amplifier with a linear voltage gain of 200 (46 dB).

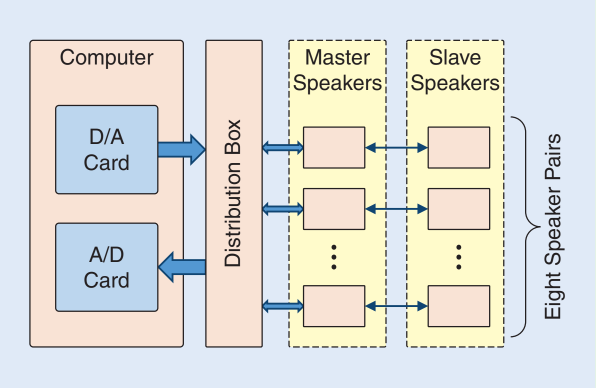


FIGURE S1. An overview of the setup.

The signal is then sent via a cable to a distribution box where signals from all speaker pairs are further amplified before they are sent to the A/D-converter card. The signals travel differentially from the speakers to the distribution box to counteract capacitively coupled noise. The last three OP-amps in the detection signal chain were placed on the collection board in this distribution box together with the corresponding circuits for the other seven speaker pairs. The first two of those OP-amps are used as voltage followers, and their main task is to remove inductively coupled noise by breaking the closed circuit that otherwise would have formed. They are also accompanied by Resistor-Capacitor-links that form first-order high-pass filters to remove any offset voltage introduced by the OP-amps in the speaker. These are the two extra OP-amps compared to a standard instrumentation amplifier. The fifth and final OP-amp in this signal chain is a differential-in single-ended-out amplifier that provides an additional linear voltage gain of 10 (20 dB).

This setup poses special demands on the OP-amps of the detection board in the master speaker box. We amplify weak signals, which means that we need low-noise amplifiers on the detection board. Also, the fairly large initial voltage gain (200) demands that the input offset voltage of these first OP-amps is low. Finally, these OP-amps must be able to operate at supply voltage 5 V, since that is what is available in the master speaker box. We use MCP6024 I/P [S6], which is a quad low-noise OP-amp that accepts supply voltage from 2.5 V to 5.5 V and delivers full output swing. Its input offset voltage is in the range ±500 μV. With voltage gain 200, this limits the output offset voltage to be in the range ±100 mV. The gain-bandwidth product of this OP-amp is 10 MHz, so with voltage gain 200, we end up with a bandwidth of 50 kHz.

The sampling frequency is determined by the A/D- and D/A-cards. They operate at a maximum of 100 ksamples/second, which is distributed between the used channels. Thus, with 16 channels, the sampling frequency used is 6,250 samples/second. The bandwidth 50 kHz of the analog part of the detection signal chain is therefore well sufficient.

While our experience from the particular choice of hardware has been positive, future generations of the

about the necessity of the documentation and the extent of the planning required, but most of them usually appreciate the LIPS model once the course is finished.

The project team normally consists of four to seven students. Generally, these students have not previously worked together, and an immediate task at the start of the project is to get to know each other and decide on the different roles that they will have in the project group. In addition to the crucial role of project manager, the group members usually assign a person responsible for documentation, quality, testing, design, and customer contact, respectively. The project task is determined in detail by the customer through the definition of a

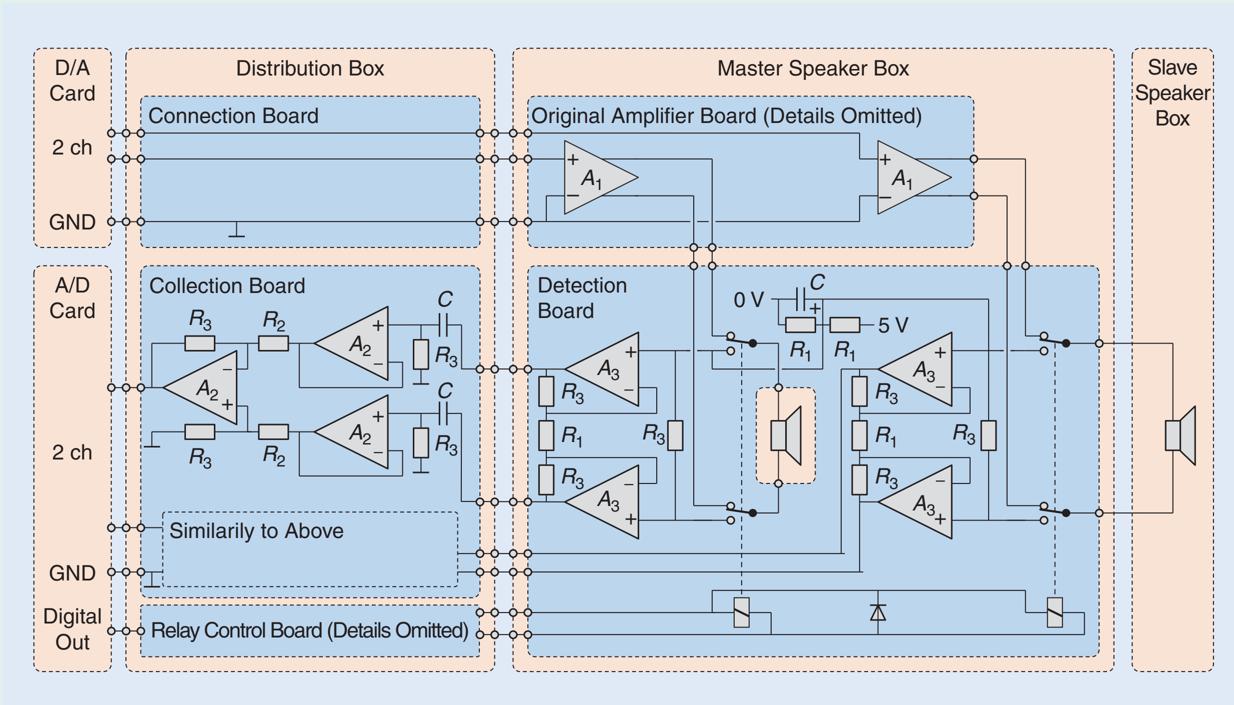


FIGURE S2. A schematic of the hardware, illustrating the circuitry for one speaker pair. A_1 is the original power amplifier for one speaker. A_2 is 1/4 of a TL074 [S7] quad OP-amp. A_3 is 1/4 of a MCP6024 [S6] quad low-noise OP-amp. Discrete component values: R_1 = 1 kΩ, R_2 = 10 kΩ, R_3 = 100 kΩ, C = 1 μF.

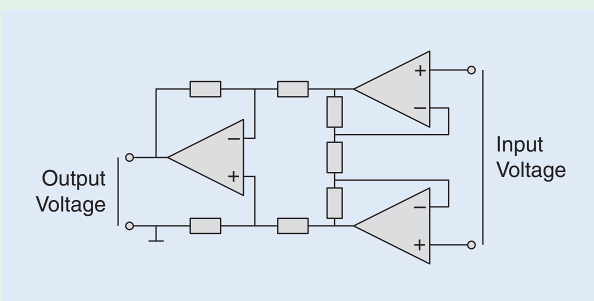


FIGURE S3. A traditional instrumentation amplifier based on three OP-amps and seven resistors.

experimental setup may use other components, in particular, eliminating the master–slave configuration inherited from the choice of the Maxxtro loudspeaker pairs, for which there is no natural need in our application.

References

[S1] Contec. Digital to analog output board for PCI, 16ch type-DA12-16(PCI), Datasheet, Ver. 2.13 [Online]. Available: http://www2.contec.co.jp/prod_data/da1216pci/da1216pci_e.pdf

[S2] Contec. (2013, Dec.). Digital to analog output board for PCI, DA12-16(PCI)-DA12-8(PCI)-DA12-4(PCI), User's Guide [Online]. Available: http://www2.contec.co.jp/dl_data/LZJ367/LZJ367_131203.pdf

[S3] Contec. Multi-channel analog input board for PCI-AD12-64(PCI), Datasheet, Ver. 3.13 [Online]. Available: http://www2.contec.co.jp/prod_data/ad1264pci/ad1264pci_e.pdf

[S4] Contec. (2013, Dec.). 64/16 Channel analog to digital input board for PCI-AD12-64(PCI)-AD12-16(PCI), User's Guide [Online]. Available: http://www2.contec.co.jp/dl_data/LZJ371/LZJ371_131202.pdf

[S5] Maxxtro. Maxxtro mini speaker MX-US-08, Datasheet [Online]. Available: http://www.elfa.spb.ru/uploads/tdpdf/kt856563.pdf

[S6] Microchip. (2009). MCP6021/1R/2/3/4, Datasheet [Online]. Available: http://ww1.microchip.com/downloads/en/DeviceDoc/21685d.pdf

[S7] Texas Instruments. (2015, June). TL07xx low-noise JFET-input operational amplifiers, Datasheet [Online]. Available: http://www.ti.com/lit/ds/symlink/tl074.pdf

set of requirements that the final product should fulfill. These tasks are usually renegotiated with the project group until a final set of requirements is agreed upon, a few weeks after the start of the project. The role of the customer is taken by a faculty member, and the course director can also assume this role. The university also provides support to the students in the form of a supervisor (normally a senior Ph.D. student or teaching assistant) as well as topic expertise (from other junior faculty members and Ph.D. students). Figure 3 depicts the various roles and the interactions, with solid lines indicating formal contacts (documents, official meetings) and dashed lines representing informal contacts (advice, meetings). Since

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

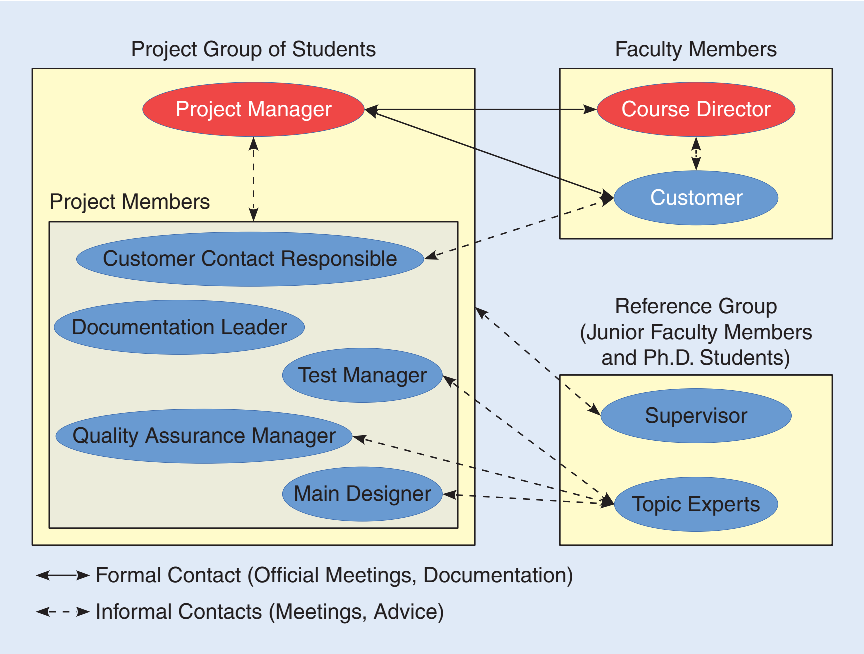


FIGURE 3. The roles of the students and faculty members in our implementation of the CDIO project course concept.

the project group members work together continuously, the everyday informal contacts between the group members are not shown in the figure.

Our general experience is that all students involved in these projects have been enthusiastic and have contributed accordingly to their tasks. The group dynamic has varied during the years and has been somewhat dependent on the interests and ambition levels of the participants. Many original ideas from students have been implemented. Usually, the desired results have been achieved. Since the time scope for the course is strictly limited to one semester, there is no room for delays. This sometimes has resulted in products that do not reach full functionality. However, in those cases the well-documented work has helped subsequent project groups to continue the development and further extend the capabilities of the product.

The purpose of our course, in general, is to provide a genuine teamwork learning experience in the field of telecommunications.

The purpose of our course, in general, is to provide a genuine teamwork learning experience in the field of telecommunications. Building an audio test bed for demonstration of the massive MIMO concept specifically was a vehicle for bringing the research frontiers to the students and has ignited their interest for the field, as well as prepared them for an industrial career. The combination of hardware and software development is a unique aspect of our course. The challenges of handling hardware imperfections has stimulated the creative skills of the involved students. Combining all of the components and implementing the desired algorithms required solid teamwork and devotion to the task. The product created so far has shown useful capabilities and is also scalable.

The course project was conducted according to a project model, which rather precisely defines rules and support for the students' work process. This helps the students organize their work in an efficient manner and make a carefully considered system design before implementing it. The project thus gives the students the opportunity to work in a fashion that is similar to how projects are conducted in industry.

In this course, we let the students work independently, without offering any initial pointers on how to solve the problem at hand. Instead, we give encouragement to invent and research potential solutions based on knowledge from previous courses, and provide access to expert supervisors who can answer questions and discuss ideas. Our general experience is that students are very successful in forming a working project group, and they enjoy the opportunity to design a system mostly on their own, learning by acquiring hands-on experience.

Lessons learned and future directions

The experimental acoustic massive MIMO setup that we have developed is inexpensive and scalable, and the learning curve for the equipment has been appropriate for the course. The project has been well-received by students in that they can apply their knowledge of communications theory to build a complete working system from scratch. They appreciate the structure of the project in that they get to try out their own ideas in solving a complex problem defined on a relatively high level. They also appreciate the access to expertise when help is needed. The students have expressed that they enjoy that the course gives them a basic understanding of the concept of massive MIMO, as an early insight into the main cutting-edge technology component of next-generation wireless networks.

Some of the students' feedback is as follows:

- We could have been more thorough in the research of equipment that was chosen for the course, since there were some compatibility issues (drivers for the A/D and D/A cards).
- The system should be expanded with more units to allow for a larger scale test system. Sixteen units is borderline "massive," and with 16 units it is difficult to get good results with four units as terminals and the remaining 12 constituting the massive MIMO array. This certainly makes for good future projects in the coming years.
- Many students would like to see how the different parts of a communication system go together in real-world systems, given more in detail in earlier courses, to be better prepared

for this kind of task. As in most universities, the majority of the regular coursework focuses on "theory," more than practical implementation aspects.

According to the students, the most important experience from the project was to create an actual, complete communication system, with real hardware. This enabled the students to use much of their knowledge from courses, while also encountering the problems that are always present with physical systems and real hardware. Solving these kinds of problems is an excellent way for the students to enhance their general skills in engineering. At the same time, they are introduced to the next-generation communication technology.

According to the students, the most important experience from the project was to create an actual, complete communication system, with real hardware.

Another aspect that the students benefit from during the project course is the opportunity to work in groups with mixed nationalities. Since the course is on the master level, typically exchange students from different countries take the course, in addition to Swedish engineering students. Thus, the students are able to work in a somewhat international environment, which is a good preparation for future employment as engineers.

Taken together, the students appreciated the opportunity to work practically with a physical system that implements the basic principles of a cutting-edge technology, massive MIMO. They also enjoy working in a very structured manner, following a project model which sets up a well-defined, logical framework for their work. This project is still relatively new, and leaves continued opportunities for improvements in the education and preparation of students for future engineering work.

Authors

Erik G. Larsson (erik.g.larsson@liu.se) received his Ph.D. degree from Uppsala University, Sweden, in 2002. He is a professor at Linköping University, Sweden. His main professional interests are within signal processing, communication theory, applied information theory, wireless systems, and 5G. He was the chair of the IEEE Signal Processing Society Signal Processing for Communications and Networking technical committee in 2015 and 2016, chair of the steering committee of *IEEE Wireless Communications Letters* in 2014 and 2015, and he organized the Asilomar Conference on Signals, Systems, and Computers (general chair 2015, technical chair 2012). He received the IEEE Signal Processing Magazine Best Column Award twice, in 2012 and 2014, and the IEEE ComSoc Stephen O. Rice Prize in Communications Theory in 2015. He is a Fellow of the IEEE.

Danyo Danev (danyo.danev@liu.se) received his M.Sc. degree in mathematics from Sofia University, Bulgaria, in 1996 and his Ph.D. degree in electrical engineering from Linköping University, Sweden, in 2001, where he is currently an associate professor teaching a number of communication engineering and mathematics courses. In 2005, he obtained the Docent title in data transmission. His research interests are within the fields of coding, information, and communication theory. He has authored or coauthored two book chapters, 16 journal papers, and more than 30 conference papers on these topics. Since 2012, he has been a board member of the IEEE Sweden Vehicular Technology/Communications/Information Theory joint Chapter.

Mikael Olofsson (mikael.olofsson@liu.se) received his M.Sc. degree in electrical engineering in 1992 and his Ph.D. degree in data transmission in 2002, both from Linköping University, Sweden. Since 2000, he has been active as a teacher at Linköping University, first as a lecturer and, since 2002, as an associate professor. He teaches electronics and communication and, in his role as a teacher, has produced a number of tutorials for his courses, of which one is a published textbook.

Simon Sörman (simon.sorman@ericsson.com) received his M.Sc. degree in 2016 in electrical engineering from Linköping University, Sweden. During his studies, he participated in the course described in this article. He is currently a researcher on wireless access networks at Ericsson Research, Linköping, and is working toward 5G development.

References

[1] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[2] E. G. Larsson, H. Q. Ngo, T. L. Marzetta, and H. Yang, Homework Problems to go with Fundamentals of Massive MIMO [1], including solution manual. [Online]. Available: www.cambridge.org/Marzetta

[3] The conceive-design-implement-operate (CDIO) [Online]. Available: http://cdio.org/

[4] J. Vieira, F. Rusek, and F. Tufvesson, "Reciprocity calibration methods for massive MIMO based on antenna coupling," in *Proc. IEEE Global Communications Conf.*, Austin, TX, 2014, pp. 3708–3712.

[5] A. Nordrum. (2016, May). "5G researchers set new world record for spectrum efficiency," *IEEE Spectr.*, [Online]. Available: http://spectrum.ieee.org/techtalk/telecom/wireless/5g-researchers-achieve-new-spectrum-efficiency-record

[6] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liu, I. Wong, V. Öwall, O. Edfors, and F. Tufvesson, "A flexible 100-antenna testbed for massive MIMO," in *Proc. IEEE GLOBECOM Workshops*, Austin, TX, 2014, pp. 287–293.

[7] TELEMIC Group. (2016). Massive MIMO 5G USRP testbed—KU Leuven [Online].Available: http://www.esat.kuleuven.be/telemic/research/NetworkedSystems/infrastructure/massive-mimo-5g

[8] P. Harris, S. Zang, A. Nix, M. Beach, S. Armour, and A. Doufexi, "A distributed massive MIMO testbed to assess real-world performance and feasibility," in *Proc. IEEE 81st Vehicular Technology Conf.*, Glasgow, Scotland, 2015, pp. 1–2.

[9] X. Yang, W. J. Lu, N. Wang, K. Nieman, S. Jin, H. Zhu, X. Mu, I. Wong, Y. Huang, and X. You, "Design and implementation of a TDD-Based 128-antenna massive MIMO prototyping system," arXiv: 1608.07362 [Online]. Available: https://arxiv.org/abs/1608.07362

[10] N. Choubey and A. Y. Panah. (2016). Introducing Facebook's new terrestrial connectivity systems—Terragraph and Project ARIES [Online]. Available: https://code.facebook.com/posts/1072680049445290/introducing-facebook-s-new-terrestrial-connectivity-systems-terragraph-and-project-aries

[11] Acoustic massive MIMO test bed (2016). [Online]. Available: https://www.youtube.com/watch?v=i1OLEEkiobA

[12] T. Svensson and C. Krysander, *Project Model LIPS*. Lund, Sweden: Student litteratur, 2011.

SP

Xiang-Gen Xia

Small Data, Mid Data, and Big Data Versus Algebra, Analysis, and Topology

Big data has become a hot topic during the last few years. But at times, its meaning has been quite confusing. I hope that through sharing my thoughts in this article, we can have a better understanding of what big data is.

Whenever you see data, you may think that it is related to numbers and counting. In fact, today, data is more general than numbers. However, when data are input in computers, they become bits and/or numbers.

So, what is big data? When was it started? Where will it lead? These questions may have different answers to different readers. To me, trained in mathematics as a signal processing professor in an electrical engineering department, data is quite natural, and so in this article, I provide my answers to the aforementioned questions.

First of all, what is big data? Unfortunately, there is no precise mathematical definition for this concept. Big data or small data is relative. To see what big data is, let us first look at what small data is. Each person in my family, which consists of four people, eats two apples per day. Therefore, my family eats eight apples per day. This is small data and is accurate. What is next? For example, my whole family, including all relatives, eats 400 apples per day. My neighbor's whole family, including all of their relatives, eats 500 apples per day. Then, the total number of apples these two families eat will be no more than 900 apples per day. You might want to ask why it is not exactly the sum, i.e., 900, of the 400 and 500 apples. The reason is that these two families may have some members in common and some of them from one family may be married to another in the other family. In this case, the total count may not be accurate, but you can have an accurate upper bound. Is this small data or something else? I would like to think of it as mid data.

Next, it comes to the number of apples consumed in the world. How many apples do the people on the earth eat per day? To find out, one might say, let us make a table of the numbers of apples eaten per day for every country. It is approximately 300 million for the United States, 300 million for Japan, etc. Oops, how many apples do the people eat in North Korea per day? Unfortunately, there is no trustworthy data available. So, what do we do? Can we still count the numbers of apples consumed per day for the whole world? No, but we may use some colors to mark the levels of the numbers for all of the countries on a map. In this case, I would consider it big data, i.e., it is so big that no one can even estimate its volume but can only get some high-level indices.

In mathematics, there are mainly three subjects: algebra, e.g., high school algebra and abstract algebra; analysis, e.g., real analysis and functional analysis; and topology and geometry, e.g., algebraic topology and differential geometry. In my opinion, all of these subjects are about counting and calculation, which is, of course, all that mathematics is about. In algebra, you can count exactly. In analysis, you may not be able to count exactly but roughly or just estimate. You might want to ask, where are probability and statistics? They belong to analysis since they belong to measure theory, which belongs to real analysis. In topology, you are not able to count the whole thing, but one still wants to count. In this case, what can be done? You can think of the whole thing as consisting of several pieces and then just count for the number of pieces. The real question is: what is a piece, and what is topology and geometry about? It is a kind of index that you may get in the limiting case. If I am asked to make an analogy between mathematics and data classification, I would say that algebra corresponds to small data, analysis corresponds to mid data, and topology/geometry corresponds to big data.

In mathematics, there are mainly three subjects: algebra, e.g., high school algebra and abstract algebra; analysis, e.g., real analysis and functional analysis; and topology and geometry, e.g., algebraic topology and differential geometry.

Small data and algebra

As discussed previously, mathematics is about counting and calculation. In

Digital Object Identifier 10.1109/MSP.2016.2607319
Date of publication: 11 January 2017

 1053-5888/17©2017IEEE

fact, calculation is a type of counting. In many calculations, finding the solutions of equations is always one of the most important tasks. Among finding the solutions of equations, finding the roots of polynomials is probably the most important. The fundamental theorem of algebra tells us that any nonconstant single-variable polynomial has at least one complex root, which means that any single-variable polynomial equation can be solved with possibly complex numbers as solutions/roots. We know that roots of a polynomial of a degree lower than five have closed forms in terms of the coefficients of the polynomial. However, for a polynomial of a degree of five or higher, its roots may not have closed forms in terms of its coefficients, which was first mathematically proven by Galois and is, therefore, called the *Galois theory*. To do so, Galois invented the concepts of group, ring, and field, which led to modern mathematics. The smallest field is the binary field {0, 1}, and the largest is the complex field C that is the set of all complex numbers. The reason why C is the largest field is because every polynomial equation over the complex field can be solved already by the fundamental theorem of algebra. There are many kinds of subfields and extended fields, such as algebraic number fields, by including, e.g., some roots of unity, i.e., $\exp(-2\pi\ j/m)$, for some positive integer *m,* in the middle of {0, 1} and C. After the complex field, mathematicians generalized C to quaternionic numbers that form, in fact, a domain as well as octonionic numbers.

For example, a quaternionic number can be equivalently written as

$$\begin{pmatrix} x & y \\ -y^* & x^* \end{pmatrix},$$

where x and y are two complex numbers. With these generalizations, mathematicians found that the most important property from all of these structures is the norm identity

$$\| x \bullet y \| = \| x \| \bullet \| y \| \qquad (1)$$

for any two elements x and y in the domain of interest, where the dot stands for the multiplication in the domain or the real multiplication, and $\| \ \|$ stands for the norm used in the domain. In other words, the norm of the product of any two elements is equal to the product of the norms of the two elements. This is clear when x and y are two complex numbers but is less obvious for other cases. A general design satisfying (1), as generalizations of complex numbers, quaternionic numbers, and octonionic numbers, is called *compositions of quadratic forms* [1]. A $[k, n, p]$ Hermitian composition formula is

$$(|x_1|^2 + \cdots + |x_k|^2)(|y_1|^2 + \cdots + |y_n|^2) \\ = |z_1|^2 + \cdots + |z_p|^2 \qquad (2)$$

where | | stands for the absolute value, $X = (x_1, \dots, x_k)$ and $Y = (y_1, \dots, y_n)$ are systems of indeterminates, and $z_i = z_i(X, Y)$ is a bilinear form of X and Y. As an example, let $k = n = p = 2$ and $z_1 = x_1 y_1 - x_2 y_2$, $z_2 = x_1 y_2 + x_2 y_1$. This corresponds to the following case. The product of the absolute values of two complex numbers is equal to the absolute value of the product of the two complex numbers, i.e., if $x = x_1 + jx_2$ and $y = y_1 + jy_2$ for real-valued x_1, x_2, y_1, y_2 and $z = z_1 + jz_2 = xy$, then $|z| = |xy| = |x\|y|$. More designs on the compositions of quadratic forms can be found in [2], which has found applications as space-time coding in wireless communications with multiple transmit antennas.

With this in mind, I would say that algebra is with the norm identity, where you are able to count precisely (the same as the first apple example mentioned previously), where $|2 \cdot 4| = 8 = |2| \cdot |4|$ and $|500 + 400| = |500| + |400|$, when the dot sign in (1) is the real multiplication and the real addition, respectively. This, in my opinion, corresponds to small data.

Mid data and analysis

In most cases, the norm identity (1) does not hold. Instead, it is the following inequality:

$$\| x \bullet y \| \leq \| x \| \bullet \| y \| \qquad (3)$$

for any two elements x and y in a set called *space*. This leads to the concept of a norm space, i.e., if there is an operation $\| \ \|$ on a set that satisfies (3) for any two elements x and y in the set, this set with some additional scaling property is called a *norm space*. It is the key for functional analysis or analysis, including measure theory and/or probability theory and statistics. In this case, in (3), the dot sign is the addition $+$, and (3) is correspondingly called the *triangular inequality*. In my opinion, the difference between algebra and analysis is the difference between the norm equality and the norm inequality shown in (1) and (3), respectively. It is the same as the second apple example mentioned previously, where

$$\| \{400 \text{ apples in one family}\} \\ \cup \{500 \text{ apples in another family}\} \| \\ \leq \| \{400 \text{ apples in one family}\} \| \\ + \| \{500 \text{ apples in another family}\} \| \\ = 400 + 500 = 900,$$

where the dot sign in (3) corresponds to the union of two sets and the real addition, respectively. I feel that this corresponds to mid data.

Another observation about the above norm inequality is that the dot operation in (3) for two elements x and y can be thought of as a general operation as we have seen above for different cases of the dot sign. The norm inequality (3) becomes the triangular inequality when the dot is $+$, as mentioned previously. When the dot is a true product of two elements, such as the matrix multiplication of two matrices, the inequality (3) is the conventional norm inequality. The norm inequality (3) becomes the Cauchy–Schwarz inequality when the dot is the inner product

$$\left| \int_a^b f(t) g(t)\, dt \right| \\ \leq \left[\int_a^b |f(t)|^2 dt \right]^{1/2} \left[\int_a^b |g(t)|^2 dt \right]^{1/2}, \qquad (4)$$

where the equality holds if, and only if, functions $f(t)$ and $g(t)$ are linearly dependent, i.e., $f(t) = cg(t)$ or $g(t) = cf(t)$ for some constant c. From this observation, almost all inequalities can be derived from the norm inequality

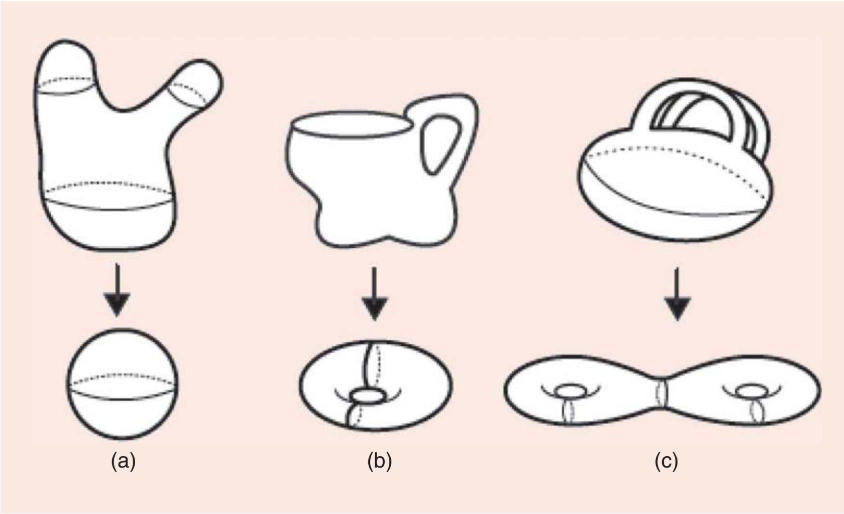


FIGURE 1. The genus of an object; (a) genus 0, (b) genus 1, and (c) genus 2.

(3). Many fundamental results are derived by the Cauchy–Schwarz inequality (4), i.e., the norm inequality. For example, the Cauchy–Schwarz inequality leads to the conclusion that the optimal linear time-invariant filter to maximize the output signal-to-noise ratio is, and only is, the filter that matches to the signal, i.e., the matched filter. It has been extensively used in radar and communications. Another application of the Cauchy–Schwarz inequality is the proof of the Heisenberg uncertainty principle (HUP). It says that the product of the time width and the bandwidth is lower bounded by one half, and the lower bound is reached if, and only if, the signal is Gaussian, i.e., $a\exp(-bt^2)$ for some constant a and some positive constant b. As a simple consequence of the HUP, one is not able to design a signal that has an infinitely small time width and infinitely small bandwidth simultaneously. Otherwise, a person would be able to design as many orthogonal signals as possible in any finitely limited area of time and frequency, i.e., it would have infinite capacity for communications over any finite bandwidth channel. One can see that both results have played key roles in science and engineering in recent history.

Big data and topology

When a person sees several large groups of fish moving in the ocean (see http://cir.institute/collective-intelligence) or large groups of birds flying in the sky (see http://becausebirds.com/2014/07/29/how-do-bird-flocks-work), he or she may not be able to count exactly or estimate approximately how many fish or birds are there. One may just count how many disconnected groups of fish. If a person treats each group as a visible hole of the ocean, it is the concept of genus, i.e., one of the key concepts in topology, where the number of holes (or fish groups in this case) in an object (i.e., the ocean) is the genus of the object. More precisely, the genus of a connected, orientable surface is an integer representing the maximum number of cuttings along nonintersecting, closed simple curves without rendering the resultant manifold disconnected [4]. In the aforementioned definition, cutting is understood as the conventional cutting by a knife. Some simple examples are shown in Figure 1. Another simple, but more mathematical, way to understand it is as follows. If any loop (i.e., a simple closed curve) on a surface (a solid object, such as a solid ball), such as the sphere shown in Figure 1(a), can be continuously (on the surface or inside the solid object) contracted/tightened (also called *continuously transformed*) to a point on the surface, then the surface has genus 0. For the torus shown in Figure 1(b), it is impossible to do so because, if one picks up a simple loop around the hole, this loop cannot be continuously contracted to any point on the surface. However, if the torus is cut in the middle with one cut, as shown in Figure 1(b) [note that there are two cuts total shown in Figure 1(b)], then it is not possible to have a loop around any hole; thus, any simple loop can be continuously contracted on the surface to a point. In this case, the torus has genus 1, i.e., one and only one cut is used/needed to do so. As shown in Figure 1, genus is a topologically invariant variable in the sense that two shapes may look totally different, but they have the same genus, where the objects in the first row have zero, one, and two holes, and are topologically equivalent to those in the second row, respectively.

> **What is big data? There is no precise mathematical definition for this concept. Big data or small data is relative.**

A possible application of the aforementioned concept of genus in topology would be in the current investigations of big data representation that plays an important role in big data analysis. One efficient way to represent big data is to use a proper tensor [5]. When big data is too big and its tensor representation is properly used, it may be treated as a multidimensional massive object. In this case, its topological properties, such as genus, may become simple but is an important feature.

As we have discussed previously, when an object is too complicated or too massive, the indices and/or the topologically invariant variables such as the genus, i.e., the number of holes and/or disconnected pieces, come to the picture. These topologically invariant variables may be obtained by taking a limit when some parameters go to infinity, which may smooth out all the uncertainties or unknowns caused by the massiveness and make the calculations possible. In other words, taking a limit may simplify the calculation. One simple example is the calculation of the integration of a Gaussian function. For any finite real values a and b and a positive constant α, $\int_a^b e^{-\alpha t^2} dt$ does not have a simple closed form while $\int_{-\infty}^{\infty} e^{-\alpha t^2} dt$ does. Another example is the diversity and multiplexing tradeoff (DMT) obtained by Zheng and Tse [3] for multiple-input, multiple-output (MIMO) antenna systems in wireless communications, which becomes a necessary parameter in designing a MIMO wireless communication system. Let R be the transmission rate in bits/second/hertz. Let r be the normalized rate $r = R/\log(\text{SNR})$, where SNR stands for signal-to-noise ratio and is the channel SNR. When SNR is huge, one may expect that R is huge as well by Shannon's channel capacity formula that is about $\log(\text{SNR})$, i.e., massive data (or big data) can be transmitted through the channel. In this case, counting R may be not possible, while counting r becomes more reasonable, where r is called the *multiplexing gain*. Let P_e be the error probability at the receiver of a MIMO modulation scheme with transmission rate R. Let

$$d(r) = -\lim_{\text{SNR}\to\infty} \frac{\log(P_e)}{\log(\text{SNR})}. \quad (5)$$

Then, $d(r)$ is the index of the negative exponential of the error probability P_e and called the *diversity gain*.

$$P_e \approx \text{SNR}^{-d(r)}, \quad (6)$$

when SNR is large enough. Zheng and Tse [3] obtained the following well-known DMT:

$$d(r) = (m-r)(n-r),$$

where m and n are the numbers of transmit and receive antennas, respectively. One can see that both r and $d(r)$ are sort of indices, and they are only meaningful when SNR approaches infinity, i.e., in a massive transmission rate case or big data case. This is the case when it is impossible to count one element by one element for a massive data, and one needs to sort out its index, such as exponentials and/or genus, in some way to describe and/or extract features from the massive/big data. I think this belongs to topology in mathematics. Thus, in my opinion, topology in mathematics corresponds to big data, where it is impossible or not necessary to count one element by one element.

I consider that small data corresponds to algebra, mid data corresponds to analysis, and big data corresponds to topology in mathematics.

Summary and discussion

In summary, I consider that small data corresponds to algebra, mid data corresponds to analysis, and big data corresponds to topology in mathematics. Was big data started when it was named? Of course not. Big data has existed for a long time, as massive groups of fish move in the ocean, massive groups of birds fly in the sky, and/or a massive number of people on the ground travel around the world. Today, massive bits are transmitted through both wired and wireless channels called the *Internet*. The key is how to get some indices, trends, or patterns from these massive data and/or how to find a needle in the ocean. What will big data lead to tomorrow? Or, how deep can we go toward infinity tomorrow? Or, how fast will a computer be tomorrow?

Author

Xiang-Gen Xia (xianggen@gmail.com) is the Charles Black Evans Professor in the Department of Electrical and Computer Engineering, University of Delaware, Newark. His main research interests include wireless communications and radar signal processing. He is a Fellow of the IEEE.

References

[1] D. B. Shapiro, *Compositions of Quadratic Forms*. New York: De Gruyter, 2000.

[2] K. Lu, S. Fu, and X.-G. Xia, "Closed-form designs of complex orthogonal space-time block codes of rate (k+1)/(2k) for 2k-1 or 2k transmit antennas," *IEEE Trans. Inform. Theory,* vol. 51, pp. 4340–4347, Oct. 2005.

[3] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073–1096, May 2003.

[4] Wikipedia. Genus. (2016). [Online]. Available: https://en.wikipedia.org/wiki/Genus_(mathematics)

[5] A. Cichocki. (2014). Era of big data processing: A new approach via tensor networks and tensor decompositions. arXiv: 1403.2048v4 (cs.ET) [Online]. Available: http://arxiv.org/abs/1403.2048v4

SP

ERRATA

Reference [8] in the "SP Education" column of the November 2016 issue of *IEEE Signal Processing Magazine* [1] was published missing a URL. We apologize for any confusion this may have caused. The corrected reference is shown in [2].

Digital Object Identifier 10.1109/MSP.2016.2636258
Date of publication: 11 January 2017

References

[1] E. Richter and A. Nehorai, "Enriching the undergraduate program with research projects," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 123–127, Nov. 2016.

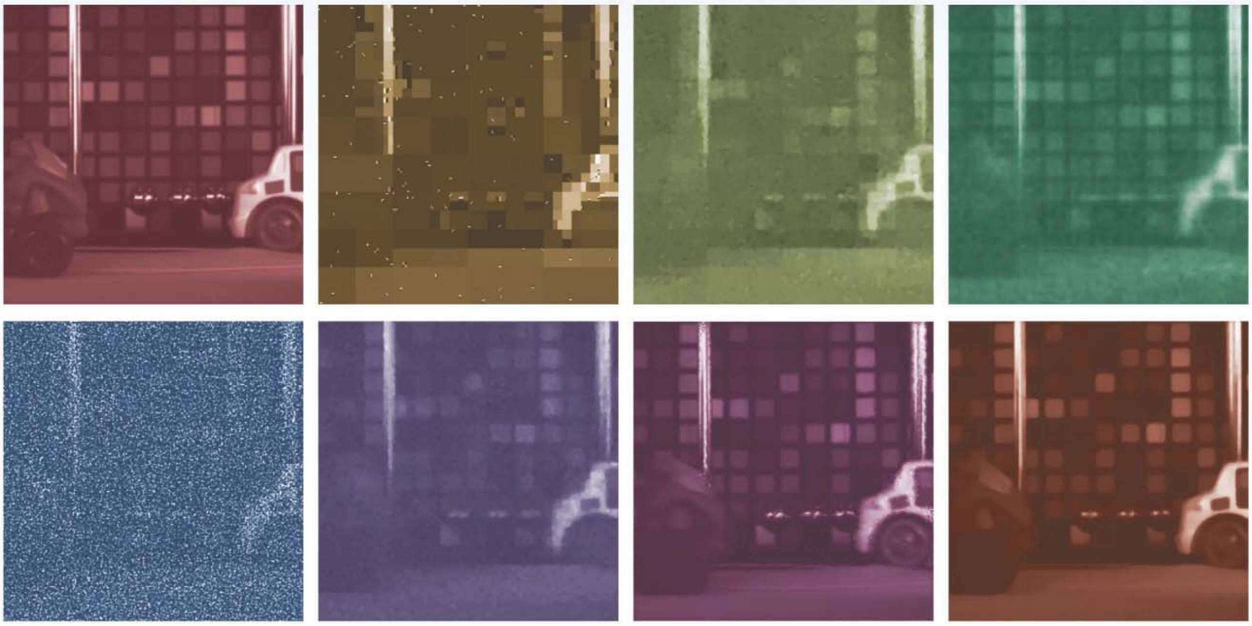
[2] BCI hand control download [Online]. Available: http://classes.engineering.wustl.edu/ese497/BigFiles/IpsiHand_v5.wmv

SP

Richard G. Baraniuk, Tom Goldstein,
Aswin C. Sankaranarayanan, Christoph Studer,
Ashok Veeraraghavan, and Michael B. Wakin

Compressive Video Sensing

Algorithms, architectures, and applications



The design of conventional sensors is based primarily on the Shannon–Nyquist sampling theorem, which states that a signal of bandwidth W Hz is fully determined by its discrete time samples provided the sampling rate exceeds 2 W samples per second. For discrete time signals, the Shannon–Nyquist theorem has a very simple interpretation: the number of data samples must be at least as large as the dimensionality of the signal being sampled and recovered. This important result enables signal processing in the discrete time domain without any loss of information. However, in an increasing number of applications, the Shannon–Nyquist sampling theorem dictates an unnecessary and often prohibitively high sampling rate (see "What Is the Nyquist Rate of a Video Signal?"). As a motivating example, the high resolution of the image sensor hardware in modern cameras reflects the large amount of data sensed to capture an image. A 10-megapixel camera, in effect, takes 10 million measurements of the scene. Yet, almost immediately after acquisition, redundancies in the image are exploited to compress the acquired data significantly, often at compression ratios of 100:1 for visualization and even higher for detection and classification tasks. This example suggests immense wastage in the overall design of conventional cameras.

Compressive sensing (CS) (see "CS 101" and [6], [14], [16], and [24]) is a powerful sensing paradigm that seeks to alleviate the daunting sampling rate requirements imposed by the Shannon–Nyquist principle. CS exploits the inherent structure (or redundancy) within the acquired signal to enable sampling and reconstruction at sub-Nyquist rates. The signal structure most commonly associated with CS is that of sparsity in a transform basis. This is the same structure exploited by image compression algorithms, which transform images into a basis [e.g., using a wavelet or discrete cosine transform (DCT)] where they are (approximately) sparse. In a typical scenario, a CS still-image camera takes a small number of coded, linear measurements of the scene—far fewer measurements than the

Digital Object Identifier 10.1109/MSP.2016.2602099
Date of publication: 11 January 2017

 1053-5888/17©2017IEEE

number of pixels being reconstructed. Given these measurements, an image is recovered by searching for the image that is sparsest in some transform basis (wavelets, DCT, or other) while being consistent with the measurements.

In essence, CS provides a framework to sense signals with far fewer measurements than their ambient dimensionality (i.e., Nyquist rate), which translates to practical benefits including decreased sensor cost, bandwidth, and time of acquisition. These benefits are most compelling for imaging modalities where sensing is expensive; examples include imaging in the nonvisible spectrum (where sensors are costly), imaging at high spatial and temporal resolutions (where the high bandwidth of sensed data requires costly electronics), and medical imaging (where the time of acquisition translates to costs or where existing equipment is too slow to acquire certain dynamic events). In this context, architectures like the single-pixel camera (SPC) [27] provide a promising proof of concept that still images can be acquired using a small number of coded measurements with inexpensive sensors.

There are numerous applications where it is desirable to extend the CS imaging framework beyond still images to incorporate video. After all, motion is ubiquitous in the real world, and capturing the dynamics of a scene requires us to go beyond static images. A hidden benefit of video is that it offers tremendous opportunities for more dramatic undersampling (the ratio of signal dimensionality to measurement dimensionality). That is, we can exploit the rich temporal redundancies in a video to reconstruct frames from far fewer measurements than is possible with still images. Yet the demands of video CS in terms of the complexity of imaging architectures, signal models, and reconstruction algorithms are significantly greater than those of compressive still-frame imaging.

There are three major reasons that the design and implementation of CS video systems are significantly more difficult than those of CS still-imaging systems. The first challenge is the gap between compression and CS. State-of-the-art video models rely on two powerful ideas: first, motion fields enable the accurate prediction of image frames by propagating intensities across frames; second, motion fields are inherently more compressible than the video itself. This observation has led to today's state-of-the-art video compression algorithms (not to be confused with CS of videos) that exploit motion information in one of many ways, including block-based motion estimation (MPEG-1), per-pixel optical flow (H.265), and wavelet lifting (LIMAT). Motion fields enable models that can be tuned to the specific video that is being sensed/processed. This is a powerful premise that typically provides an order of magnitude improvement in video compression over image compression.

The use of motion fields for video CS raises an important challenge. Unlike the standard video compression problem,

What Is the Nyquist Rate of a Video Signal?

Conventional videos, sampled at 24–60 frames/second (fps), may, in fact, be highly undersampled in time—objects in the scene can move multiple pixels between adjacent frames. Some compressive sensing (CS) architectures, however, measure a video at a much higher temporal rate. For example, the single-pixel camera (SPC) may take tens of thousands of serial measurements per second. In such cases, the scene may change very little between adjacent measurements. This raises some interesting questions: what is the Nyquist rate of a video signal, and how does it compare to CS measurement rates?

One can gain insight into these questions by considering the three-dimensional analog video signal that arrives at a camera lens; both conventional and CS imaging systems can be viewed as blurring this signal spatially (due to the optics and the pixelated sensors) and sampling or measuring it digitally. If a video consists of moving objects with sharp edges, then the analog video will actually have infinite bandwidth in both the spatial and temporal dimensions. However, it can be argued that the support of the video's spectrum will tend to be localized into a certain bowtie shape, as shown in blue in Figure S1. The salient feature of this shape is that high temporal frequencies coincide only with high spatial frequencies. Thus, because of the limited spatial resolution of

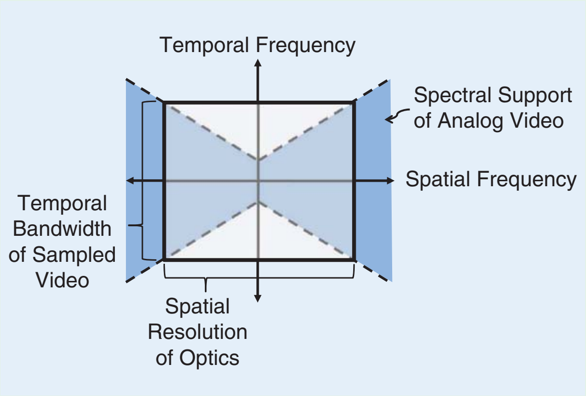


FIGURE S1. The limited spatial resolution of an imaging system may also limit its temporal bandwidth.

both the camera optics and the pixel sensors, when the spatial bandwidth of the video is limited, so too is its temporal bandwidth, as illustrated by the black rectangle in the figure. This suggests that the video sensed by architectures such as the SPC may in fact have a finite temporal bandwidth, and this fact can be used to reduce the computational complexity of sensing and reconstructing the video. In particular, it is not necessary to reconstruct at a rate of thousands of fps. Additional details are provided in [62].

CS 101

Compressive sensing (CS) exploits the fact that a small and carefully selected set of nonadaptive linear measurements of a compressible signal, image, or video carries enough information for reconstruction and processing [16], [24]; for a tutorial treatment see [6], [14].

The traditional digital data acquisition approach uniformly samples the three-dimensional analog signal corresponding to the time variations of a scene; the resulting samples $V[x, y, t]$ in space (x, y) and time (t) are sufficient to perfectly recover a bandlimited approximation to the scene at the Nyquist rate. Let the abstract vector s represent the Nyquist-rate samples of the scene $V[x, y, t]$; see "What Is the Nyquist Rate of a Video Signal?" for a discussion of the Nyquist rate of a time-varying scene. Because the number of samples required for real-world scenes, N, is often very large, for example, in the billions for today's consumer digital video cameras, the raw image data is typically reduced via data compression methods that typically rely on transform coding.

As an alternative, CS bypasses the Nyquist sampling process and directly acquires a compressed signal representation using $M < N$ linear measurements between s and a collection of linear codes $\{\phi[m]\}_{m=1}^{M}$ as in $y[m] = \langle s, \phi[m] \rangle$. Stacking the measurements $y[m]$ into the M-dimensional vector y and the transpose of the codes $\phi[m]^T$ as rows into an $M \times N$ sensing matrix Φ, we can write $y = \Phi s$.

The transformation from s to y is a dimensionality reduction and does not, in general, preserve information. In particular, because $M < N$, there are infinitely many vectors s' that satisfy $y = \Phi s'$. The magic of CS is that Φ can be designed such that sparse or compressible signals s can be recovered exactly or approximately from the measurements y. By *sparse,* we mean that only $K \ll N$ of the entries in s are nonzero, or that there exists a sparsifying transform Ψ such that most of the coefficients of $\alpha := \Psi s$ are zero. By *compressible,* we mean that s or α is approximately sparse. Let $\Psi^{-1} := [\psi_1, \psi_2, \ldots, \psi_N]$ represent the inverse of the $N \times N$ basis matrix; then, $s = \Psi^{-1}\alpha$ and $y = \Phi s = \Phi\Psi^{-1}\alpha$.

Typically in CS, the sparse signal s or its sparse coefficients α is recovered by solving an optimization problem of the form (1), where f measures the fidelity of the recovery (e.g., using the squared error $\| y - \Phi\Psi^{-1}\alpha \|_2^2$) and g is a regularization penalty (e.g., the ℓ_1-norm $\|\alpha\|_1$, which promotes sparsity of α). In these cases, the resulting problem is convex, which guarantees a single global minimizer that can be found using a range of algorithms.

While the design of the sensing matrix Φ is beyond the scope of this review, typical CS approaches employ a random matrix. For example, we can draw the entries of Φ as independent and identically distributed ± 1 random variables from a uniform Bernoulli distribution [8]. Then, the measurements y are merely M different sign-permuted linear combinations of the elements of s. Other choices for Φ exist in the literature, such as randomly subsampled Fourier or Hadamard bases. In this case, multiplication by Φ can be accomplished using fast transform algorithms, which enables faster reconstruction than is possible with random matrices.

It is important to emphasize that CS is not a panacea for all the world's sampling problems [7]. In particular, to apply the concept profitably, it is critical that the signal s possess a lower inherent dimensionality than its ambient dimensionality (e.g., sparse structure) and that the degree of undersampling N/M be balanced with respect to the signal's signal-to-noise ratio [22].

where the frames of the video are explicitly available to perform motion estimation, in CS, we have access only to coded and undersampled measurements of the video. We are thus faced with a chicken-or-egg problem. Given high-quality video frames, we could precisely estimate the motion fields; but we need precise motion estimates in the first place to obtain high-quality video frames. The second challenge is the laws of causality and imaging architectures. Time waits for no one. A distinguishing property of the video sensing problem over still imaging is the fundamental difference between space and time. The ephemeral nature of time poses significant limitations on the measurement process—clearly, we cannot obtain additional measurements of an event after it has occurred. As a consequence, it is entirely possible that a compressive camera does not capture a sufficient number of measurements to recover the frames of the video. Overcoming this challenge requires both an understanding of the spatial-temporal resolution tradeoffs associated with video CS and development of novel compressive imaging architectures that can deliver very high measurement rates or reconstruct at different resolutions depending on the available data. The third challenge is computational complexity. Even moderate resolution videos result in high bandwidth streaming measurements. Typical CS video recovery algorithms are highly nonlinear and often involve expensive iterative optimization routines. Fast (or even real-time) reconstruction of CS video is challenging because it requires a data measurement system, fast iterative algorithms, and high-performance hardware jointly designed to enable sufficiently high throughput.

The goal of this article is to overview the current approaches to video CS and demonstrate that significant gains can be obtained using carefully designed CS video architectures and algorithms. However, these gains can only be realized when there is cohesive progress across three distinct fields: video models, compressive video sensing architectures, and video reconstruction algorithms. This article reviews progress that has been made in advancing and bringing these fields together.

We discuss some of the landmark results in video CS and highlight their key properties and the rich interplay among models, architectures, and algorithms that enable them. We also lay out a research agenda to attack the key open research problems and practical challenges to be resolved in video CS.

Video sensing systems

In this section, we discuss the current compressive imaging architectures that have been proposed for CS video. The architectures can be broken down into three categories (see Table 1).

- *Spatial multiplexing cameras* (*SMCs*): SMCs optically superresolve a low-resolution sensor to boost spatial resolution. SMCs are invaluable in regimes where high-resolution sensors are unavailable, as in terahertz/millimeter-wave and magnetic resonance imaging (MRI), or extremely costly, as in short or medium wavelength infrared (SWIR and MWIR) sensing.
- *Temporal multiplexing cameras* (*TMCs*): TMCs optically superresolve a low-frame-rate camera to boost temporal resolution. TMCs are mainly used to overcome the limitations imposed on the measurement rate by the analog-to-digital converter (ADC) and are optimized to produce a high-frame-rate video at high spatial resolution with low-frame-rate sensors.
- *Spectral and angular multiplexing cameras* (*SAMCs*): SAMCs boost resolution in the spectral domain, which can be useful for hyperspectral and light-field video sensing. As with TMCs, the bottleneck of these architectures is also the measurement rate constraint imposed by the ADC.

Each of these flavors of a CS system aims to break the Nyquist barrier to obtain either higher spatial, temporal, or spectral resolution. In the following sections we discuss the key design considerations and existing implementations of these three camera types.

SMCs

SMCs apply CS multiplexing in space to boost the spatial resolution of images and videos obtained from sensor arrays with low spatial resolution. The use of a low-resolution sensor enables SMCs to operate at wavelengths where corresponding full-frame sensors are too expensive, such as at SWIR, MWIR, terahertz, and millimeter wavelengths. SMCs employ a spatial light modulator, such as a digital micromirror device (DMD) or liquid crystal on silicon (LCoS), to optically compute a series of coded inner products with the rasterized scene s; these linear inner products determine the rows of the sensing matrix Φ (recall the notation from "CS 101"). It is worth mentioning that the SMC approach

Table 1. The key architectures for CS video and their properties.

Type	Name	Application	Modulator	Best-known capabilities	Limitations
SMC	SPC	Infrared imaging	DMD	Spatial resolution 128 × 128 Time resolution 64 fps Result [27]	Operational speed of DMD
	LiSens/FPA-CS	Infrared imaging	DMD	Spatial resolution 1,024 × 768 Time resolution 10 fps Result [19], [78]	Need for precise optical alignment/calibration
TMCs	Coded strobing	High-speed imaging	Mechanical/ferroelectric shutter	Spatial resolution (sensor) Time resolution 2,000 fps Result [75]	Periodic scenes
	Flutter shutter	High-speed imaging	Mechanical/ferroelectric shutter	Spatial resolution (sensor) Time resolution 4 × sensor fps Result [64]	Locally linear motion
	P2C2	High-speed imaging	LCoS	Spatial resolution (sensor) Time resolution 16 × sensor fps Result [65]	Dynamic range of sensor
	Per-pixel shutter	High-speed imaging	LCoS/electronic shutter	Spatial resolution (sensor) Time resolution 16 × sensor fps Result [39]	Light loss
	CACTI	High-speed imaging	Translating mask	Spatial resolution (sensor) Time resolution 100 × sensor fps Result [51]	Mechanical motion
Light-field video		Dynamic refocusing	LCoS, used as programmable coded aperture	Time resolution sensor fps Result [71]	Loss of spatial resolution can be severe for high spectral/angular resolutions
Hyperspectral video	CASSI	Spectroscopy	Static mask	Time resolution sensor fps Result [76]	

fps: frames/second; FPA: focal plane array; P2C2: programmable pixel compressive camera; CACTI: coded aperture compressive temporal imaging; CASSI: coded aperture snapshot spectral imaging.

is equally applicable to modalities outside the scope of this article, such as MRI [52], where the physics of image formation produces a measurement system that can be interpreted as subsampling the Fourier transform of the sensed image.

Fast (or even real-time) reconstruction of CS video is challenging because it requires a data measurement system, fast iterative algorithms, and high-performance hardware jointly designed to enable sufficiently high throughput.

SPC

The SPC [27] acquires images using only a single sensor element (i.e., a single pixel) and taking significantly fewer multiplexed measurements than the number of scene pixels. In the SPC, light from the scene is focused onto a programmable DMD, which directs light from a subset of activated micromirrors onto the single photodetector. The programmable nature of the DMD's micromirror orientation enables one to direct light either toward or away from the photodetector. As a consequence, the voltage measured at the photodetector corresponds to an inner product of the image focused on the DMD and the micromirrors directed toward the sensor (see Figure 1). Specifically, at time t, if the DMD pattern is represented by $\phi[t]$ and the time-varying scene by $V[x,y,t]$ (where x and y are the two spatial dimensions and t is the temporal dimension), then the photodetector measures a scalar value $y[t] = \langle \phi[t], V[\cdot,\cdot,t] \rangle + e[t]$, where $\langle \cdot,\cdot \rangle$ denotes the inner product between the vectors and $e[t]$ accounts for the measurement noise. If the scene is static, that is, $V[x,y,t] = V_0[x,y]$, then the measurement vectors can be stacked as columns into a measurement matrix, with $\Phi = [\phi_1, \phi_2, \ldots, \phi_M]^T$. The SPC leverages the relatively high pattern rate of the DMD, which is defined as the number of unique micromirror configurations that can be obtained in unit time. This pattern rate, typically 10–20 kHz for commercially available devices, defines the measurement bandwidth (i.e., the number of measurements per second) and is one of the key factors that defines the achievable spatial and temporal resolutions. Because SPCs rely on the DMD to modulate images onto a single sensor, the spatial resolution is limited by the density of mirrors on the DMD.

Since the proposal of the original SPC in [27], numerous authors have developed alternative SPC architectures that do not require a DMD for spatial light modulation. In [41], a liquid-crystal display panel is used for spatial light modulation; the use of a transmissive light modulator enables a lensless architecture. Sen and Darabi [70] use a camera-projector system to construct an SPC exploiting a concept referred to as *dual photography* [69]; the hallmark of this system is its use of active and coded illumination that can be beneficial in certain applications, particularly microscopy.

Beyond SPCs—Multipixel detectors

As mentioned previously, the measurement rate of an SPC is limited by the pattern rate of its DMD, which is typically in the tens of kilohertz. This measurement rate can be insufficient for scenes with very high spatial and temporal resolutions. This issue can be combatted using an SMC with F sensor pixels (photodetectors), each capturing light from a nonoverlapping region of the DMD. The measurement rate of the SMC increases linearly with the number of photodetectors. Taking into account that the maximum measurement rate is capped by the sampling rate of the ADC, we can write the measurement rate for an SMC with F photodetectors as

$$\min\{F \times R_{\text{DMD}}, R_{\text{ADC}}\},$$

where R_{DMD} is the pattern rate of the DMD and R_{ADC} is the sampling rate of the ADC. Hence, the smallest number of photodetectors for which the measurement rate is maximized is

$$\text{(minimum number of sensor pixels)} \quad F_{\min} = R_{\text{ADC}}/R_{\text{DMD}}.$$

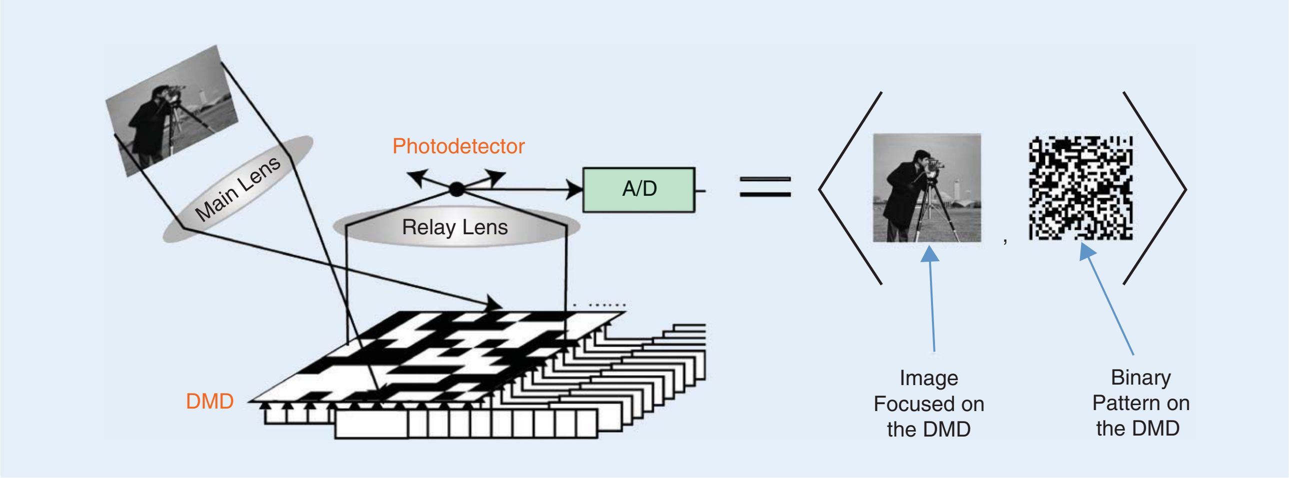


FIGURE 1. The operation principle of the SPC. Each measurement corresponds to an inner product between the binary mirror–mirror orientation pattern on the DMD and the scene to be acquired. (Figure courtesy of [67].)

In essence, at $F = F_{\min}$ we can obtain the measurement rate of a full-frame sensor but using a device with potentially a fraction of the number of photodetectors. This can be invaluable for sensing in many wavebands, for example, SWIR.

As a case study, consider an SMC with a DMD pattern rate $R_{\mathrm{DMD}} = 10$ kHz and an ADC with a sampling rate $R_{\mathrm{ADC}} = 10$ MHz. Then, for a sensor with $F_{\min} = 1{,}000$ pixels, we can acquire 10 million measurements per second. An SPC, in comparison, would acquire only 10,000 measurements per second. Consequently, multipixel SMCs can acquire videos at significantly higher spatial and temporal resolutions than an SPC.

There have been many multipixel extensions to the SPC concept. The simplest approach [46] maps the DMD to a low-resolution sensor array, as opposed to a single photodetector, such that each pixel on the sensor observes a nonoverlapping patch or a block of micromirrors on the DMD. SMCs based on this design have been proposed for sensing in the visible [78], SWIR [19], and MWIR [54]. Figure 2 shows an example of the increased measurement rates offered by the LiSens camera [78], which uses a linear array of 1,024 photodetectors. More recently, there have also emerged multipixel multiplexing-based cameras that completely get rid of the lens and replace the lens with a mask and computational reconstruction algorithms [2].

TMCs

TMCs apply CS multiplexing in time to boost the temporal resolution of videos obtained from sensor arrays with low temporal resolution. Again, let $V[x,y,t]$ be a three-dimensional (3-D) signal representing a time-varying scene. Due to the assumed low frame rate of the sensor, we obtain a scene measurement once every T seconds, where T is too large. If the SLM has an operational speed of one pattern every T_{SLM} seconds, then each measurement of a TMC takes the form of a coded image:

$$y[x,y,t_0] = \sum_{j=0}^{C-1} \phi[x,y,j]\, V[x,y,t_0 + jT_{\mathrm{SLM}}],$$

where $\phi[x,y,j]$ is the attenuation pattern on the SLM at spatial location (x, y) and time jT_{SLM}. Here, each coded image measured by the TMC multiplexes C frames of the high-speed

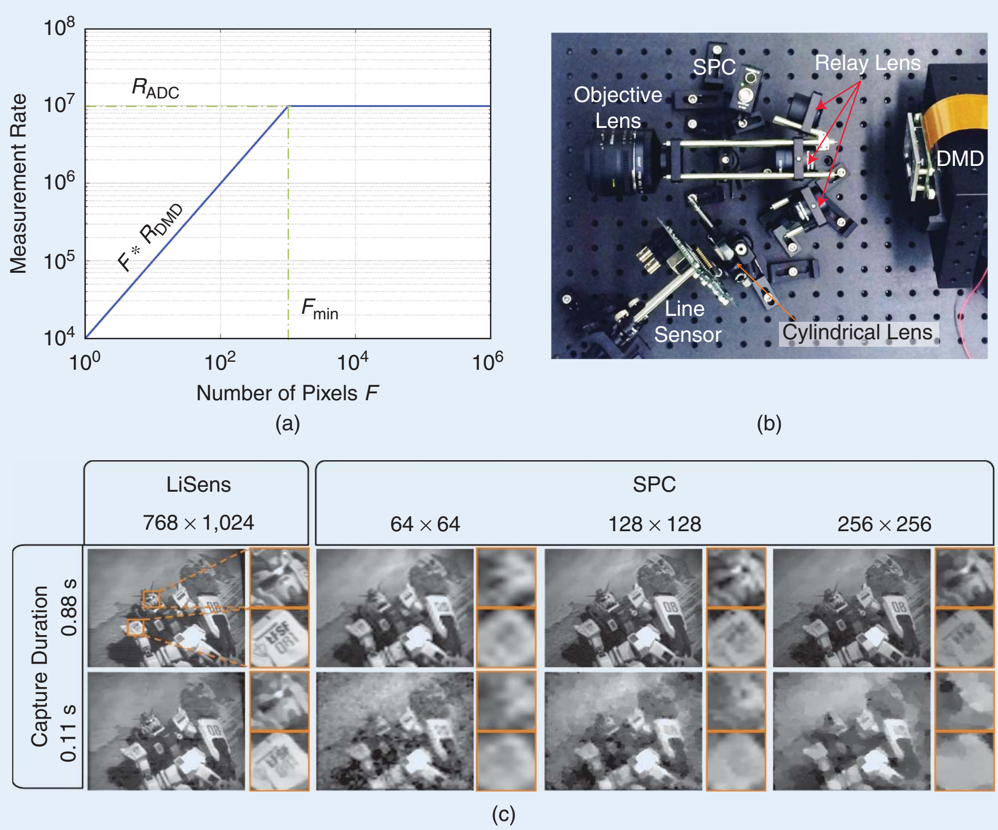


FIGURE 2. The multipixel SMCs support significantly higher sensing rates than an SPC. (a) The measurement rate as a function of the number of sensor pixels. An optimized SMC with $F_{\min}$ pixels delivers the highest possible measurement rate. (b) Lab prototypes of the SPC and LiSens cameras, each placed on the one arm of a single DMD. The measurement rate of the LiSens camera is nearly 1 MHz, while that of the SPC is 20 kHz. (c) Comparisons between LiSens, which uses 1,024 sensor pixels, and an SPC for a static scene. Each row corresponds to a different capture duration, defined as the total amount of time that the cameras have for acquiring compressive measurements. The larger measurement rate of the LiSens camera enables it to sense scenes with very high spatial resolution even for small capture durations. (Photos courtesy of [78].)

video, and hence, we obtain one coded image every CT_{SLM} seconds. Our goal is to recover the frames of the high-speed video $V[x,y,kT_{\mathrm{SLM}}]$ from a single or a sequence of coded images/measurements.

SMCs apply CS multiplexing in space to boost the spatial resolution of images and videos obtained from sensor arrays with low spatial resolution.

Global shutters
The simplest instance of a TMC uses a global shutter together with a conventional camera. In a global shutter, the SLM code $\Phi[x,y,j] = \Phi[j]$ is spatially invariant, which can be implemented by using a programmable shutter or by using the image sensor's built-in electronic shutter. Veeraraghavan et al. [75] showed that periodic scenes can be imaged at very high temporal resolutions using a global shutter [64]. This idea has been extended to nonperiodic scenes in [40], where a union-of-subspace model was used to temporally superresolve the captured scene. However, global shutters are fundamentally limited to providing only spatially invariant coding of the video; this can be insufficient to provide a rich-enough encoding of a high-speed video. Hence, in spite of their simplicity, global shutters fail for scenes with complex motion patterns.

Per-pixel shutters
Reddy et al. [65] proposed the programmable pixel compressive camera (P2C2), which extends the global shutter idea with per-pixel shuttering. Here, each pixel has its own unique code that is typically binary valued and pseudorandom. The P2C2 architecture uses an LCoS SLM placed optically at the sensor plane and carefully aligned to a high-resolution two-dimensional (2-D) sensor. The P2C2 prototype achieves $16\times$ temporal superresolution, even for complex motion patterns. Hitomi et al. [39] extended the P2C2 camera using a per-pixel coding that is more amenable to implementation in modern image sensors with per-pixel electronic shutters. Here, $\Phi[x,y,j] = \delta[j - j_0(x,y)]$; that is, each pixel observes the intensity at one of the subframes of the high-speed video, and the selection of this subframe varies spatially. Llull et al. [51] and Koller et al. [47] proposed a TMC that achieves temporal multiplexing via a translating mask in the sensor plane. This approach avoids the hardware complexity involved with DMD and LCoS SLMs and enjoys other benefits, including low operational power consumption at the cost of having a mechanical component (the translating mask).

Additional TMC designs
Gu et al. [36] used the rolling shutter of a complementary metal–oxide–semiconductor (CMOS) sensor to enable higher temporal resolution. The key idea is to stagger the exposures of each row randomly and use image/video models to recover a high-frame-rate video. Harmany et al. [37] extended coded aperture systems by incorporating a global shutter; the resulting TMC provides immense flexibility in the choice of the measurement matrix Φ.

SAMCs
SAMCs apply CS multiplexing to sense variations of light in a scene beyond the spatial and temporal dimensions. Two specific examples include hyperspectral CS video cameras that sense spatial, spectral, and temporal variations of light in a scene and light-field video cameras that sense spatial, angular, and temporal variations. In both cases, imaging at high resolution across all modalities simultaneously requires that we handle both high measurement rates (this is typically limited by the ADC sampling rate) and low light levels (due to scene light being resolved into various modalities). CS techniques, more specifically, signal models, can address both bottlenecks. Examples of compressive cameras include the coded aperture snapshot spectral imaging architecture [76] and compressive hyperspectral imaging using spectrometers [50] for spectral multiplexing and the work of Marwah et al. [58] and Tambe et al. [71] for angular multiplexing.

Models for video structure
Recovering a video from compressive linear measurements requires one to extract the video signal s from the measurements $y = \Phi s$ (recall "CS 101"). Here, s might represent a certain block of pixels, an entire video frame, or an ensemble of frames, depending on the sensing architecture and the specific recovery algorithm employed. All of these are functions of the underlying time-varying scene $V[x,y,t]$. Because the number of measurements M is less than the video signal's ambient dimensionality N, infinitely many vectors s' may satisfy $y = \Phi s'$. Hence, to recover s from y, a model that captures the scene structure (or a priori information) of s with a small number of degrees of freedom is required; the model can then be included in the recovery algorithm. This section surveys several popular models for characterizing low-dimensional structure in videos.

Single-frame structure
The structure of a single video frame can be characterized using standard models for conventional 2-D images. Natural images have been shown to exhibit sparse representations in the 2-D DCT, 2-D wavelet, and curvelet domains [15], [56]. Images have also been shown to have sparse gradients. The total variation (TV) seminorm promotes such gradient sparsity simply by minimizing the ℓ_1 norm of an image's 2-D gradient [52]. To fully exploit the structure in a 3-D video, one needs to characterize the spatial and temporal dimensions simultaneously, rather than reconstructing each frame independently and only accounting for spatial structure. Hence, the spatial 2-D regularizers described previously often appear as building blocks of more sophisticated 3-D video models.

Sparse innovation models
One of the simplest possible models accounting for multi-frame structure assumes that a video can be reduced into a static and a dynamic component. This model—while restrictive—is applicable, for example, in surveillance applications,

where a scene is observed from a distant static camera. We can decompose each frame of such a video into a static background frame and a number of small (sparse) foreground objects that may change location from frame to frame. A natural way of modeling such structure is to assume that the differences between consecutive frames have a sparse representation in some transform basis. That is, for two consecutive video frames $V[x,y,t_1]$ and $V[x,y,t_2]$, one may assume that the difference frame $V[x,y,t_2]-V[x,y,t_1]$ has a sparse representation in a basis such as a 2-D wavelet basis. Such models have been explored in detail in the context of CS [17], [57], [74] and can be viewed as special cases of the more advanced motion-compensation techniques described below.

One of the simplest possible models accounting for multiframe structure assumes that a video can be reduced into a static and a dynamic component.

Low-rank matrix models

An alternative approach to scene modeling involves reorganizing a 3-D video signal into a 2-D matrix, where each column of the matrix contains a rasterized ordering of the pixels of one video frame. A variety of popular concise models for matrix structure can then be interpreted as models for video structure. One of the most prominent models asserts that the matrix is low rank; this is equivalent to assuming that the columns of the data matrix live in a common, low-dimensional subspace. In the context of video modeling, a seminal result by Basri and Jacobs [9] showed that collections of images of a Lambertian object under varying lighting often cluster close to a nine-dimensional subspace. This property can be useful for modeling videos of stationary scenes where the illumination conditions change over time.

To account for both variations in background illumination and for sparse foreground objects that move with time, one can extend the low-rank matrix model to a low-rank-plus-sparse model [79], [80]. A sparse matrix, added to the original low-rank matrix, accounts for sparse foreground innovations, such as small moving objects. Again, such models are particularly suitable for surveillance applications.

TV minimization and sparse dictionaries

Sparsifying transforms such as wavelets, curvelets, and the DCT have natural extensions to 3-D [56], [77], [82] and can be employed for jointly reconstructing an ensemble of video frames. TV minimization can also be extended to 3-D [35], [49]; minimizing the 3-D TV seminorm of a video promotes frames with sparse gradients across spatial and temporal dimensions.

It is also possible to learn specialized (possibly overcomplete) bases that enable sparse representations of patches, frames, and videos from training data. A variety of so-called dictionary learning algorithms have been proposed that learn sparsifying frames Ψ (see, e.g., [1] and "CS 101"). Dictionary learning algorithms can be used not only to generate dictionaries that sparsify images but also to sparsify videos in both the spatial and temporal dimensions. This approach has been successfully employed for CS video reconstruction in [39].

Linear dynamical systems

Linear dynamical systems (LDSs) model the dynamics in a video using linear subspace models. Such models have been used extensively in the context of activity analysis and dynamic textures. Video CS using LDS reduces to the estimation of the LDS parameters, including the observation matrix and the state transition matrix, from compressive measurements. Approaches for parameter estimation have included recursive [73] as well as batch methods [66]. Furthermore, [66] demonstrates the use of the recovered LDS parameters for activity classification.

Motion compensation

While regularizers such as 3-D wavelets and 3-D TV minimization can be used for CS video reconstruction, it is worth noting that conventional video compression algorithms (such as H.264) do not employ such simple techniques. Rather, because objects in a video may move (or translate) several pixels between adjacent frames, it is typical to employ block-based motion compensation and prediction, where each video frame is partitioned into blocks, the location of each block is predicted in the next frame, and only the residual of this prediction is encoded.

Some CS video architectures may require reconstructions of video sequences with high temporal frame rates. In these cases, there may be relatively little object motion between consecutive frames. Consequently, motion compensation may not be required, and techniques such as 3-D TV may result in high-quality scene recovery.

In other cases, however, it may be necessary to predict and compensate for the motion of objects between consecutive frames. This presents an interesting chicken-or-egg problem: motion compensation can help in reconstructing a video, but the motion predictions themselves cannot be made until (at least part of) the video is reconstructed. One iterative, multiscale technique has been proposed [62] that alternates between motion estimation and video reconstruction: the recovered video at coarse scales (low spatial resolution) is used to estimate motion, which is then used to boost the recovery at finer scales (high spatial resolution). Given the estimated motion vectors, a motion-compensated 3-D wavelet transform can be defined using the LIMAT technique [68]. Another approach initially reconstructs frames individually, estimates the motion between the frames, and then attempts to reconstruct any residual not accounted for by the motion prediction [30]; see also [45] for a related technique. The logistics of block-based video sensing and reconstruction are discussed in detail in [30].

Optical flow

A more general approach to motion compensation involves the optical flow field. Given two frames of a video, $V[x,y,t_1]$ and

$V[x, y, t_2]$, optical flow refers to the flow field $\{u(x,y), v(x,y)\}$ such that $V[x+u(x,y), y+v(x,y), t_1] = V[x,y,t_2]$. Optical flow enables one to represent the frames of a video using a small collection of key frames plus optical flow fields that synthesize (extrapolate) the video from the key frames. Optical flow fields are often significantly more compressible than images. Such an approach is closely related to the block-based motion compensation models described earlier but is distinguished by its explicit attempt to model motion on a per-pixel basis.

A key challenge in the use of optical flow models for video CS is—once again—that, in the context of sensing, we do not have access to the flow fields nor do we have access to high-quality images from which to estimate the flow fields. Reddy et al. [65] resolve this chicken-or-egg problem by first recovering a video with simple image-based priors, estimating the optical flow field on the initial reconstruction, and subsequently recovering the video again while simultaneously enforcing the brightness constancy constraints derived using optical flow. They show that a 30-frames/second (fps) sensor can be superresolved to a 240–480-fps sensor by temporal modulation using an LCoS device. In the context of SMCs, Sankaranarayanan et al. [67] use a specialized dual-scale sensing (DSS) matrix that provides robust and computationally inexpensive initial scene estimates at a lower spatial resolution. This enables this approach to robustly estimate optical flow fields on a low-resolution video. Optical flow-based video CS has also been applied for the dynamic MRI problem, where carefully selected Fourier measurements provide robust initial scene estimates [3]. The concept of DSS sensing matrices has been improved recently by the sum-to-one (STOne) transform [35], which enables the fast recovery of low-resolution scene estimates at multiple resolutions.

Video recovery techniques

While the mathematical formulations of video CS recovery problems resemble other canonical sparse recovery problems, three important factors set video recovery apart from other types of sparse coding. First, video recovery problems are extremely large and have high memory requirements. Methods for high-resolution video recovery must scale to hundreds of millions of unknowns. Second, sparse representations of videos with complex structures may contain tens of thousands (or more) of nonzero entries. Consequently, algorithm implementations that require large dense matrix systems are intractable, and methods must exploit fast transforms. Third, high-quality video recovery often involves noninvertible sparsity transforms, and so reconstruction methods that handle cosparsity models are desirable. Some recovery problems require more sophisticated (or unstructured) models, such as optical flow constraints, that cannot be handled efficiently by simple algorithms. All of these factors impact algorithm performance on different reconstruction applications.

This section overviews the range of existing recovery techniques and investigates the tradeoffs between reconstruction quality and computational complexity. For simplicity, we focus on two categories of reconstruction methods, variational and greedy. Note that there are algorithms that do not fit well into these categories (such as iterative hard thresholding [12], which has features of both); a discussion of such methods is beyond the scope of this article.

Variational methods

Variational methods for CS video recovery perform scene reconstruction by solving optimization problems using iterative algorithms. Most variational methods suitable for high-dimensional problems can be classified into two categories, constrained and unconstrained, as detailed next.

Constrained problems

The first category solves constrained problems of the form

$$\hat{s} = \arg\min_{s,z} f(\Phi s \mid y) + g(z) \quad \text{subject to } z = \Psi s. \tag{1}$$

Here, the function f models the video acquisition process (optics, modulation, and sampling), and g is a regularizer that promotes sparsity under the transformation defined by Ψ. For example, basic frame-by-frame recovery with 2-D wavelet sparsity can be formulated as an unconstrained problem with $f(\Phi s \mid y) = \| y - \Phi s \|_2^2$ and $g(z) = \gamma \| z \|_1$, where s contains a vectorized image frame, Φ is the sensing matrix, Ψ is a 2-D wavelet transform, and $\gamma > 0$ is a regularization parameter. Under a TV scene model, the matrix Ψ is a discrete gradient operator that computes differences between adjacent pixels. 3-D TV video recovery can be achieved by stacking multiple vectorized video frames into s and defining Ψ to be the 3-D discrete gradient across both spatial dimensions and time. Optical flow constraints can be included by forming a sparse matrix Ψ that differences pixels in one frame with pixels that lie along its flow trajectory in other frames.

It can be shown that the solution to (1) corresponds to a saddle point of the so-called augmented Lagrangian function

$$\mathcal{L}(s,z,\lambda) = f(\Phi s \mid y) + g(z) + \frac{\beta}{2}\| z - \Psi s - \lambda \|_2^2, \tag{2}$$

where λ is a vector of Lagrange multipliers. Constrained problems of the form (1) for CS video can be solved efficiently using the alternating direction method of multipliers (ADMM) [13], [28], [31] or the primal-dual hybrid gradient (PDHG) method [18], [29]. The ADMM and PDHG methods alternate between minimization steps for s and z and maximization steps for λ until convergence is reached. Such methods have the key advantage that they enable the inclusion of powerful, noninvertible video models such as 3-D TV or optical flow. This advantage, however, comes at the cost of higher memory requirements and somewhat more complicated iterations. To improve the convergence rates of solvers for constrained problems, accelerated algorithm variants have been developed [18], [32], [33].

Unconstrained problems

If the sparsity transform Ψ is invertible, then the constraint in (1) can be removed by replacing the vector s with $\Psi^{-1}z$. This leads to the second category of recovery methods that solve unconstrained problems of the following simpler form:

$$\hat{z} = \arg\min_{z} f(\hat{\Phi} z | \ y) + g(z). \tag{3}$$

Here, the matrix $\hat{\Phi} = \Phi\Psi^{-1}$ and z contains the representation of a single frame or the entire video in the sparsity transform domain. For example, in the case of wavelet sparsity, solving (3) recovers the video's wavelet coefficients; the final video is obtained by applying the inverse wavelet transform to the solution.

Unconstrained problems of the form (3) can be solved efficiently using forward–backward splitting (FBS) [20], fast iterative shrinkage/thresholding (FISTA) [10], fast adaptive shrinkage/thresholding algorithm (FASTA) [34], sparse reconstruction by separable approximation (SpaRSA) [81], or approximate message passing (AMP) [25], [55]. FBS is the most basic variant for solving unconstrained problems and performs the following two steps for the iterations $k = 1, 2, \ldots$ until reaching convergence:

$$\hat{z}^{k+1} = z^k + \tau^k \hat{\Phi}^* \nabla f(\hat{\Phi} z^k | y) \text{ and} \tag{4}$$

$$z^{k+1} = \arg\min_{z} g(z) + \frac{1}{2} \| z - \hat{z}^{k+1} \|_2^2, \tag{5}$$

where $\{\tau^k\}$ is some step size sequence. FBS finds a global minimum of the objective function (3) by alternating between the explicit gradient-descent step (4) in the function f and the proximal (or implicit gradient) step (5) in the function g. The key operations of the gradient step (4) are matrix–vector multiplications with $\hat{\Phi}$ and $\hat{\Phi}^*$. These multiplications can be carried out efficiently when $\hat{\Phi}$ is a composition of fast transforms, such as subsampled Hadamard/Fourier matrices and wavelet or DCT operators. When g is a simple sparsity-promoting regularizer, such as the ℓ_1 norm, the proximal step (5) is easy to compute in closed form using wavelet shrinkage. The computational complexity of FBS can be reduced significantly using adaptive step-size rules for selecting $\{\tau^k\}$, acceleration schemes, restart rules, momentum (or memory) terms, and so forth, as is the case for FISTA, FASTA, SpaRSA, and AMP. See the review article [34] for more details.

Greedy pursuit algorithms

Greedy pursuit algorithms are generally used for unconstrained problems and iteratively construct a sparse set of nonzero transform coefficients. Each iteration begins by identifying a candidate sparsity pattern for the unknown vector z. Then, a least-squares problem is solved to minimize $\| \hat{\Phi} z - y \|_2^2$, where z is constrained to have the prescribed sparsity pattern.

Existing greedy pursuit algorithms can be classified into sequential greedy pursuit algorithms and parallel greedy pursuit algorithms. Sequential methods include orthogonal matching pursuit (OMP), regularized OMP (ROMP), and stagewise OMP (StOMP) [26], [61], [72]. These methods successively add more and more indices to the support set until a maximum sparsity K is reached. Parallel methods, such as compressive sampling matching pursuit (CoSaMP) and subspace pursuit [21], [60], constantly maintain a full support set of K nonzero entries but add strong and replace weak entries in an iterative fashion. Parallel greedy pursuit algorithms have the advantage that they can enforce structured models on the support set, such as a wavelet tree structure [5].

The main drawbacks of greedy algorithms, however, are that 1) they are typically unable to handle noninvertible sparsity transforms used for video reconstruction such as TV, optical flow, or overcomplete wavelet frames; 2) accurate solutions are guaranteed only when the measurement operator satisfies stringent conditions (such as the restricted isometry property or similar incoherence conditions [60], [72]); and 3) they require solving large linear systems on every iteration. For small numbers of unknowns (<10,000), the factorization of these systems can be explicitly represented and updated cheaply using rank-one updates. For the large video CS problems considered here, iterative (conjugate gradient) methods are recommended. These methods require only matrix multiplications (which can exploit fast transforms) and have lower memory requirements because they do not require the storage of large and dense matrices.

Reconstruction quality versus computational complexity

There are many choices to make when building a compressive video pipeline, including measurement operators, video models, and reconstruction algorithms. Most reconstruction algorithms are restricted as to what measurement operators and sparsity models they can support. To achieve the best performance, the reconstruction algorithms, video models, and data acquisition pipelines must be designed jointly; this implies that there are tradeoffs to be made among reconstruction speed, algorithm simplicity, and video quality.

The classical approach to CS video recovery is to search for the video that is compatible with the observed measurements while being as sparse as possible in the wavelet domain. When an invertible wavelet transform is used, the reconstruction problem can be transformed into an unconstrained problem of the form (3), which can be solved efficiently using variational methods such as FBS. If we further assume that the wavelet transform is orthogonal, then we can use off-the-shelf greedy pursuit algorithms, such as CoSaMP. Unfortunately, while unconstrained optimization is simple to implement and highly efficient, wavelet-based scene priors generally result in lower reconstruction quality than noninvertible/redundant sparsity models like TV. For this reason, we are often interested in constrained solvers that interface with TV-based video models and optical flow constraints.

To examine the associated performance/complexity tradeoffs, we compare a variety of reconstruction methods using the same measurement operator. A stream of 65,536 STOne measurements [35] was acquired from a 256 × 256 pixel video having 16 frames. Videos were reconstructed separately using various models and solvers that were implemented in MATLAB. We consider unconstrained recovery using CoSaMP and FBS, which are restricted to using invertible regularizers. In the wavelet case, we consider 1) 2-D frame-by-frame

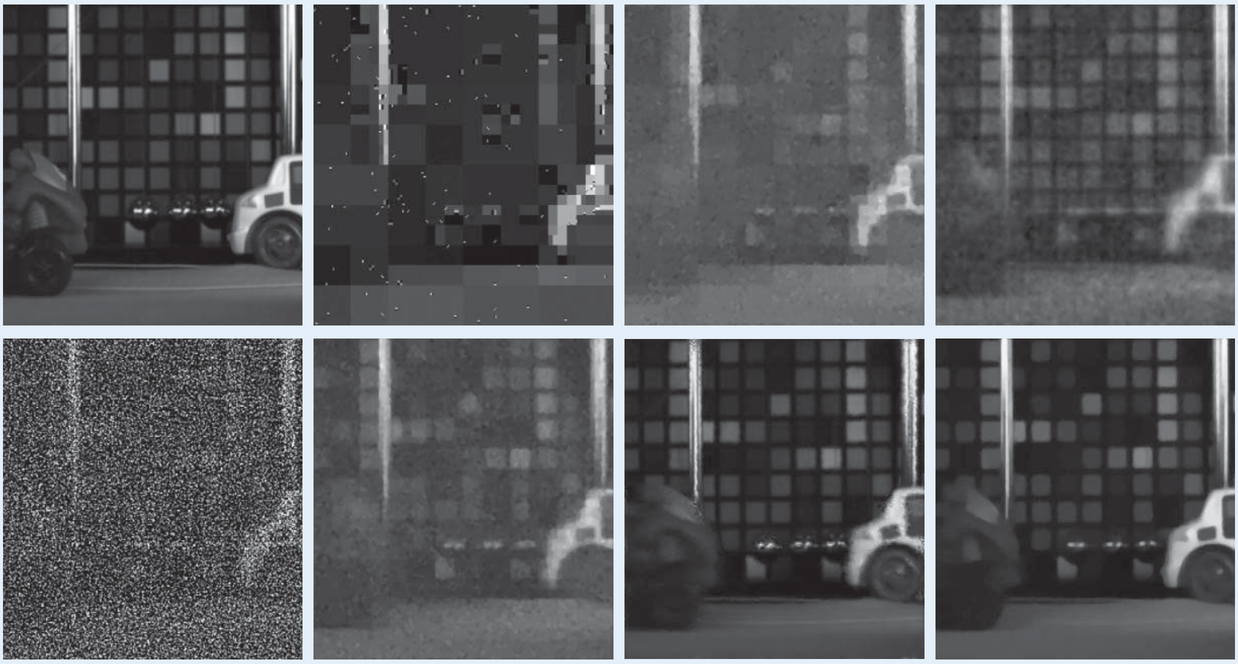


FIGURE 3. A CS video recovery comparison with different video models. For each model, we recover a 16-frame video with 256×256 pixel resolution from 2^{16} STOne transform measurements, corresponding to a 16:1 compression ratio. Sparsity models include 2-D (across space) and 3-D (across space and time) wavelet sparsity using the Haar wavelet, the 3-D DCT, optical flow constraints, and 3-D TV. For each experiment, we also provide the total runtime for recovering 16 frames.

recovery that does not exploit correlations across time, and 2) 3-D wavelet recovery that performs a 3-D wavelet transform across space and time. We also consider sparsity under the 3-D DCT, which is invertible and enjoys extensive use in image and video compression. We furthermore consider solvers for constrained problems that handle more sophisticated sparsity models. In particular, we compare 3-D TV models with PDHG and optical flow constraints with ADMM (as in CS-MUVI [67]). As a baseline, we perform CS video recovery without scene priors by simply computing $\Phi^T y$, the product of the adjoint of the measurement operator with the vector of measurements. Because the measurement operator is a subsampled orthogonal matrix, this corresponds to a least-squares recovery using the pseudoinverse. All experiments are carried out on an off-the-shelf laptop with 16 GB memory and a 2.6 GHz i5 central processing unit (CPU) with two physical cores (no parallelism was used for reconstruction).

Sample frames from our experiments together with the required runtime are shown in Figure 3. We observe that TV regularization and optical flow models dramatically outperform wavelet-based recovery in terms of video quality. Furthermore, 3-D models lead to significantly improved image quality with fewer artifacts than 2-D models, despite the fact that both reconstructions see the same amount of data. This demonstrates the efficacy of exploiting correlations across time. The key advantage of 2-D models is that they enable parallel frame-by-frame reconstruction, for example, by dispatching different recovery problems on separate CPU cores. Finally, we see that for these types of large-scale reconstruction problems, variational methods require substantially lower runtimes than greedy pursuit algorithms. The CoSaMP result in Figure 3 is for frame-by-frame reconstructions with a sparsity level of $K = 256$ nonzero wavelet coefficients per image. CoSaMP's runtime increases dramatically for larger K or when 3-D regularizers are used. This is because each iteration requires the solution to a large least-squares problem using multiple iterative (conjugate gradient) steps. Hence, such greedy pursuit algorithms turn out to be efficient only for highly sparse signals and not for general CS video problems.

Perspectives and open research questions

The video CS problem has spawned a growing body of research that spans signal representations and models, computational sensing architectures, and efficient optimization techniques. This has led to a vibrant ecosystem of methodologies that have transitioned the theoretical ideas of CS into concrete application-specific concepts. We conclude by highlighting some of the important open questions and future research directions.

IEEE SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

Real-time CS video recovery with today's hardware

High-quality CS video recovery requires complex algorithms that include powerful video models. While offline video recovery is always feasible, reconstruction using more sophisticated scene models (e.g., using optical flow) can easily take several seconds to minutes even for only a few low-resolution frames. As a consequence, applications that necessitate real-time video recovery face extreme implementation challenges. From our experiments in Figure 3, we see that even the fastest algorithms with basic video models are more than $20\times$ to $200\times$ below real time when executed in MATLAB on off-the-shelf CPUs.

Quite surprisingly, when counting the number of floating-point operations (FLOPs) required for the main transforms of these methods, we observe that real-time CS video recovery with variational methods is within reach of existing hardware. In fact, variational-based scene recovery of a 256×256 pixel scene at 12 fps requires only about 20 GFLOPs, which is well below that of programmable processing hardware, such as CPUs, graphics-processing units (GPUs), and field-programmable gate arrays (FPGAs) that achieve peak throughputs of a few TFLOPs. Similarly, existing application specific integrated circuit (ASIC) designs that target CS recovery problems [11], [53] are able to solve variational problems with more than 200 GOPS (the computations are typically carried out with fixed-point arithmetic instead of floating point) using low silicon area and low power when implemented in modern CMOS technology nodes. In Figure 4, we compare the complexity versus the resolution of various CS video recovery methods. One can observe that even higher resolutions like 1080p HD are feasible in real time with computationally efficient algorithms. Nevertheless, no real-time CS video recovery implementation has been proposed in the open literature, which can mainly be attributed to the lack of highly optimized and massively parallel CS video recovery pipelines for programmable hardware (CPUs, GPUs, or FPGAs) as well as dedicated integrated circuits (ASICs). This is definitely a fruitful area for future work.

Compressive inference rather than recovery

The main results of CS are directed toward providing novel sampling theorems that determine the feasibility of signal reconstruction from an underdetermined set of linear measurements. However, reconstruction is often not the eventual goal in most applications, which range from detection and classification to tracking and parameter estimation. While these tasks can all be performed postreconstruction (on the output of a reconstruction procedure), there are important benefits to be gained by performing them directly on the compressive measurements. First, tasks like detection, classification, and tracking are inherently simpler than reconstruction—hence, there is hope that we can perform them with fewer measurements. Second, CS reconstruction is intrinsically tied to the signal models used for the unknown signal, and these signal models prioritize features that deal with visual perception, which often is not the most relevant for the subsequent processing tasks.

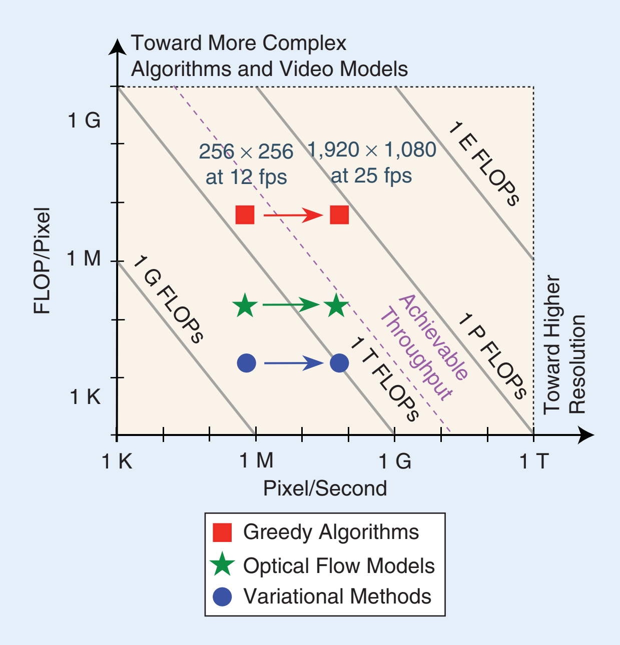


FIGURE 4. The complexity (in FLOPs per pixel) versus resolution (in pixels per second) for greedy algorithms, variational methods, and optical-flow models for the video scene in Figure 3. Variational methods (including 3-D TV and 3-D/2-D wavelets) require the lowest complexity and enable real-time CS video recovery with existing hardware (the diagonal dotted line shows the FLOPs limit of current reprogrammable hardware). Optical flow models exceed the capabilities of current hardware and require the development of more efficient computational methods and faster processing architectures.

Third, as previously discussed at length, CS reconstruction algorithms have high computational complexity, and hence avoiding a reconstruction step in the overall processing pipeline can be beneficial.

There has been some limited work on inference from linear compressive measurements. Davenport et al. [23] perform compressive classification and detection by using a matched filter in the compressive domain. Their key observation is that random projections preserve distances as well as inner products between sparse vectors; thus, inference tasks like hypothesis testing and certain filtering operations can be performed directly in the compressive domain. Hegde et al. [38] show that manifold learning (or nonlinear dimensionality reduction) can be performed just as well on the compressive measurements as on the original data, provided the data arises from a manifold with certain smoothness properties. Sankaranarayanan et al. [66] demonstrate that for time-varying systems well approximated as linear dynamical systems, the parameters of the dynamical system can be directly estimated given compressive measurements. Recently, Kulkarni and Turaga [44] proposed a novel method based on recurrence textures for action recognition from compressive cameras especially for self-similar feature sequences [43]. Apart from these early attempts, there is very little in the literature exploring high-level inference from compressive imagers.

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

A major hurdle to successful compressive inference in the video context is the mismatch between part-based models, used in computer vision, and global random embeddings, the cornerstone of the CS theory. Part-based models have had remarkable success over the past decade in object detection and classification problems. The key enabler of part-based inference is a local feature description that helps isolate objects from background clutter and provides robustness against object variations. However, the conventional CS measurements are dense random projections that are not conducive to local feature extraction without reconstructing the signal first. Hence, there is an urgent need for CS measurement operator designs that enable local feature extraction.

From measurements to bits—Toward nonlinear sensing architectures

One of the important distinctions between video CS and video compression is the nature of representing the compressed data. Compression aims to reduce the number of bits used to represent the video. In contrast, CS measurements are typically represented in terms of real values with infinite (or arbitrarily large) precision; here, the number of actual measurements is the criterion to reduce/optimize. The focus on reducing the number of measurements is often misplaced in many sensing scenarios; for example, in high-speed video CS, the bottleneck is solely due to the operating speed of the ADC, whose performance is measured in the number of bits acquired per second. Hence, compressively sensing while respecting the bottlenecks imposed by the ADC sampling frequency requires us to consider measurements in terms of bits. While there has been some effort in the area of 1-bit CS [4], [42], [63] and the tradeoff between measurement bits and measurement rate [48], this aspect is still largely unexplored in literature. In particular, there is a need for new kinds of nonlinear sensing architectures that optimize system performance in the context of the practical realities of sensing (quantization, saturation, etc.). Some initial progress in this direction for CS has been made in [59], but the area remains wide open for research.

Acknowledgments

We thank David Robert Jones for his invaluable suggestions and Doug Jones for the JAM. Richard G. Baraniuk was supported by National Science Foundation (NSF) grants CCF-1527501 and CCF-1502875, Defense Advanced Research Projects Agency (DARPA) Revolutionary Enhancement of Visibility by Exploiting Active Light-fields grant HR0011-16-C-0028, and Office of Naval Research (ONR) grant N00014-15-1-2735. Tom Goldstein was supported by NSF grant CCF-1535902 and ONR grant N00014-15-1-2676. Aswin C. Sankaranarayanan was supported by NSF grant IIS-1618823 and Army Research Office grant W911NF-16-1-0441. Christoph Studer was supported in part by Xilinx Inc. and by NSF grants ECCS-1408006 and CCF-1535897. Ashok Veeraraghavan was supported by NSF grant CCF-1527501. Michael B. Wakin was supported by NSF CAREER grant CCF-1149225 and grant CCF-1409258.

Authors

Richard G. Baraniuk (richb@rice.edu) received his B.S. degree from the University of Manitoba, Canada, in 1987; his M.S. degree from the University of Wisconsin-Madison in 1988; and his Ph.D. degree from the University of Illinois at Urbana-Champaign in 1992. He is the Victor E. Cameron Professor of Electrical and Computer Engineering at Rice University, Houston, Texas, and the founder and director of OpenStax (openstax.org). His research interests include new theory, algorithms, and hardware for sensing, signal processing, and machine learning. He is a Fellow of the IEEE and the American Association for the Advancement of Science and has received national young investigator awards from the National Science Foundation and the Office of Naval Research; the Rosenbaum Fellowship from the Isaac Newton Institute of Cambridge University, United Kingdom; the Electrical and Computer Engineering Young Alumni Achievement Award from the University of Illinois at Urbana–Champaign; the IEEE Signal Processing Society Best Paper, Best Column, Education, and Technical Achievement Awards; and the IEEE James H. Mulligan, Jr. Medal.

Tom Goldstein (tomg@cs.umd.edu) received his B.A. degree in mathematics from Washington University, St. Louis, Missouri, in 2006 and his Ph.D. degree in mathematics from the University of California, Los Angeles, in 2010. He has been a visiting research scientist with Stanford University, California, and Rice University, Houston, Texas. He is currently an assistant professor of computer science at the University of Maryland, College Park. His research interests include numerical optimization, distributed computing, image processing, and machine learning.

Aswin C. Sankaranarayanan (saswin@andrew.cmu.edu) received his B.S. degree in electrical engineering from the Indian Institute of Technology, Madras, in 2003 and his Ph.D. degree from the University of Maryland, College Park, where he was awarded the distinguished dissertation fellowship by the Department of Electrical and Computer Engineering (ECE) in 2009. He was a postdoctoral researcher in the Digital Signal Processing group at Rice University, Houston, Texas. He is currently an assistant professor in the ECE Department at Carnegie Mellon University, Pittsburgh, Pennsylvania. His research encompasses problems in compressive sensing and computational imaging. He has received best paper awards at the Computer Vision and Pattern Recognition Workshops on Computational Cameras and Displays (2015) and Analysis and Modeling of Faces and Gestures (2010).

Christoph Studer (studer@cornell.edu) received his M.S. and Ph.D. degrees from ETH Zurich, Switzerland, in 2005 and 2009, respectively. He has been a postdoctoral student and research scientist at ETH Zurich and Rice University, Houston, Texas, and is currently an assistant professor in the School of Electrical and Computer Engineering, Cornell University, Ithaca, New York. His research interests are at the intersection of digital VLSI circuit and system design, signal and image processing, and wireless communication.

Ashok Veeraraghavan (vashok@rice.edu) received his B.S. degree in electrical engineering from the Indian Institute of Technology, Madras, in 2002 and his M.S. and Ph.D. degrees from the Department of Electrical and Computer Engineering, University of Maryland, College Park, in 2004 and 2008, respectively. He is an assistant professor of electrical and computer engineering at Rice University, Houston, Texas, where he directs the Computational Imaging and Vision Lab. His research interests are broadly in the areas of computational imaging, computer vision, and robotics. Before joining Rice University, he spent three years as a research scientist at Mitsubishi Electric Research Labs in Cambridge, Massachusetts. His work has received numerous awards, including the doctoral dissertation award from the Department of Electrical and Computer Engineering at the University of Maryland, the Hershel M. Rich Invention Award from Rice University, and the Best Poster Runner-Up Award from the 2014 International Conference on Computational Photography.

Michael B. Wakin (mwakin@mines.edu) received his B.S. degree in electrical engineering and his B.A. degree in mathematics in 2000, his M.S. degree in electrical engineering in 2002, and his Ph.D. degree in electrical engineering in 2007 from Rice University, Houston, Texas. He is the Ben L. Fryrear associate professor in the Department of Electrical Engineering and Computer Science at the Colorado School of Mines (CSM), Golden. He was a National Science Foundation (NSF) mathematical sciences postdoctoral research fellow at the California Institute of Technology, Pasadena, in 2006–2007 and an assistant professor at the University of Michigan, Ann Arbor, in 2007–2008. His research interests include sparse, geometric, and manifold-based models for signal processing and compressive sensing. He has received the NSF CAREER Award, the Defense Advanced Research Projects Agency Young Faculty Award, and the CSM Excellence in Research Award for his research as a junior faculty member.

References

[1] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[2] M. S. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, and R. Baraniuk, "Flatcam: Thin, lensless cameras using coded aperture and computation," *IEEE Trans. Comput. Imag.*, to be published. DOI: 10.1109/TCI.2016.2593662

[3] M. S. Asif, L. Hamilton, M. Brummer, and J. Romberg, "Motion-adaptive spatio-temporal regularization for accelerated dynamic MRI," *Magn. Reson. Medicine,* vol. 70, no. 3, pp. 800–812, 2013.

[4] R. Baraniuk, S. Foucart, D. Needell, Y. Plan, and M. Wootters, "Exponential decay of reconstruction error from binary measurements of sparse signals," arXiv preprint arXiv:1407.8246, 2014.

[5] R. G. Baraniuk, "Optimal tree approximation with wavelets," *Proc. SPIE*, vol. 3813, pp. 196–207, 1999.

[6] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Process. Mag.,* vol. 24, no. 4, pp. 118–121, 2007.

[7] R. G. Baraniuk. (2015). Compressive nonsensing, Norbert Weiner Lecture, Univ. of Maryland. [Online]. Available: http://www.norbertwiener.umd.edu/FFT/2015/15-TAs/baraniuk.html

[8] R. G. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[9] R. Basri and D. W. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 25, no. 2, pp. 218–233, 2003.

[10] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.,* vol. 18, no. 11, pp. 2419–2434, 2009.

[11] D. E. Bellasi, L. Bettini, C. Benkeser, T. Burger, Q. Huang, and C. Studer, "VLSI design of a monolithic compressive-sensing wideband analog-to-information converter," *IEEE Trans. Emerg. Sel. Topics Circuits Syst.,* vol. 3, no. 4, pp. 552–565, 2013.

[12] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.,* vol. 27, no. 3, pp. 265–274, 2009.

[13] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[14] E. Candès and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[15] E. J. Candès and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise c2 singularities," *Commun. Pure Appl. Math.,* vol. 57, no. 2, pp. 219–266, 2004.

[16] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[17] V. Cevher, A. C. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa, "Compressive sensing for background subtraction," in *Proc. European Conf. Computer Vision*, Marseille, France, 2008, pp. 155–168.

[18] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Mathematical Imaging Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[19] H. Chen, S. Asif, A. C. Sankaranarayanan, and A. Veeraraghavan, "FPA-CS: Focal plane array-based compressive imaging in short-wave infrared," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* Boston, MA, 2015, pp. 2358–2366.

[20] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering.* New York: Springer-Verlag, 2011, pp. 185–212.

[21] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.

[22] M. Davenport, J. Laska, J. R. Treichler, and R. G. Baraniuk, "The pros and cons of compressive sensing for wideband signal acquisition: Noise folding versus dynamic range," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4628–4642, 2012.

[23] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal processing with compressive measurements," *IEEE J. Sel. Topics Signal Process.,* vol. 4, no. 2, pp. 445–460, 2010.

[24] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory,* vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[25] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 106, no. 45, pp. 18914–18919, 2009.

[26] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory,* vol. 58, no. 2, pp. 1094–1121, 2012.

[27] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.

[28] J. Eckstein and D. P. Bertsekas, "On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming,* vol. 55, nos. 1–3, pp. 293–318, 1992.

[29] E. Esser, X. Zhang, and T. F. Chan, "A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science," *SIAM J. Imag. Sci.,* vol. 3, no. 4, pp. 1015–1046, 2010.

[30] J. E. Fowler, S. Mun, E. W. Tramel, M. R. Gupta, Y. Chen, T. Wiegand, and H. Schwarz, "Block-based compressed sensing of images and video," *Found. Trends Signal Processing,* vol. 4, no. 4, pp. 297–416, 2010.

[31] R. Glowinski and P. Le Tallec, Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics. Philadelphia, PA: SIAM, 1989.

[32] T. Goldstein, M. Li, and X. Yuan, "Adaptive primal-dual splitting methods for statistical learning and image processing," in *Advances in Neural Information Processing Systems 28*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, R. Garnett, and R. Garnett, Eds. Red Hook, NY: Curran Associates, Inc., 2015, pp. 2080–2088.

[33] T. Goldstein, B. O'Donoghue, S. Setzer, and R. Baraniuk, "Fast alternating direction optimization methods," *SIAM J. Imag. Sci.,* vol. 7, no. 3, pp. 1588–1623, 2014.

[34] T. Goldstein, C. Studer, and R. Baraniuk, "A field guide to forward–backward splitting with a fasta implementation," *arXiv preprint arXiv*:1411.3406, 2014.

[35] T. Goldstein, L. Xu, K. F. Kelly, and R. G. Baraniuk, "The STOne transform: Multi-resolution image enhancement and real-time compressive video," *arXiv preprint arXiv*:1311.3405, 2013.

[36] J. Gu, Y. Hitomi, T. Mitsunaga, and S. Nayar, "Coded rolling shutter photography: Flexible space-time sampling," in *Proc. 2010 IEEE Int. Conf. Computational Photography*, Cambridge, MA, pp. 1–8.

[37] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "Compressive coded aperture keyed exposure imaging with optical flow reconstruction," *arXiv preprint arXiv*:1306.6281, 2013.

[38] C. Hegde, M. Wakin, and R. Baraniuk, "Random projections for manifold learning," in *Proc. 22nd Annu. Conf. Neural Information Processing Systems*, Vancouver, Canada, 2008, pp. 641–648.

[39] Y. Hitomi, J. Gu, M. Gupta, T. Mitsunaga, and S. K. Nayar, "Video from a single coded exposure photograph using a learned over-complete dictionary," in *Proc. 2011 Int. Conf. Computer Vision*, Barcelona, Spain, pp. 287–294.

[40] J. Holloway, A. C. Sankaranarayanan, A. Veeraraghavan, and S. Tambe, "Flutter shutter video camera for compressive sensing of videos," in *Proc. 2012 IEEE Int. Conf. Computational Photography*, Seattle, WA, pp. 1–9.

[41] G. Huang, H. Jiang, K. Matthews, and P. Wilford, "Lenless imaging by compressive sensing," in *Proc. Int. Conf. Image Processing*, Melbourne, Australia, 2013, pp. 2101–2105.

[42] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Trans. Inf. Theory,* vol. 59, no. 4, pp. 2082–2102, 2013.

[43] I. Junejo, E. Dexter, I. Laptev, and P. Perez, "View-independent action recognition from temporal self-similarities," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 33, no. 1, pp. 172 –185, 2011.

[44] K. Kulkarni and P. Turaga, "Recurrence textures for activity recognition using compressive cameras," in *Proc. 2012 19th IEEE Int. Conf. Image Processing*, Lake Buena Vista, FL, pp. 1417–1420.

[45] L.-W. Kang and C.-S. Lu, "Distributed compressive video sensing," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Taiwan, China, 2009, pp. 1169–1172.

[46] K. Kelly, R. Baraniuk, L. McMackin, R. Bridge, S. Chatterjee, and T. Weston, "Decreasing image acquisition time for compressive imaging devices," U.S. Patent 8,860,835, Oct. 14, 2014.

[47] R. Koller, L. Schmid, N. Matsuda, T. Niederberger, L. Spinoulas, O. Cossairt, G. Schuster, and A. K. Katsaggelos, "High spatio-temporal resolution video with compressed sensing," *Opt. Express,* vol. 23, no. 12, pp. 15992–16007, 2015.

[48] J. Laska and R. G. Baraniuk, "Regime change: Bit-depth versus measurement-rate in compressive sensing," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3496–3505, 2012.

[49] Y. Le Montagner, E. Angelini, and J.-C. Olivo-Marin, "Video reconstruction using compressed sensing measurements and 3d total variation regularization for bio-imaging applications," in *Proc. 2012 19th IEEE Int. Conf. Image Processing*, Lake Buena Vista, FL, pp. 917–920.

[50] C. Li, T. Sun, K. F. Kelly, and Y. Zhang, "A compressive sensing and unmixing scheme for hyperspectral data processing," *IEEE Trans. Image Process.,* vol. 21, no. 3, pp. 1200–1210, 2012.

[51] P. Llull, X. Liao, X. Yuan, J. Yang, D. Kittle, L. Carin, G. Sapiro, and D. J. Brady, "Coded aperture compressive temporal imaging," *Opt. exp., vol.* 21, no. 9, pp. 10526–10545, 2013.

[52] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Process. Mag.,* vol. 25, no. 2, pp. 72–82, 2008.

[53] P. Maechler, C. Studer, D. E. Bellasi, A. Maleki, A. Burg, N. Felber, H. Kaeslin, and R. G. Baraniuk, "VLSI design of approximate message passing for signal restoration and compressive sensing," *IEEE J. Emerg. Sel. Topics Circuits Syst.,* vol. 2, no. 3, pp. 579–590, 2012.

[54] A. Mahalanobis, R. Shilling, R. Murphy, an d R. Muise, "Recent results of medium wave infrared compressive sensing," *Appl. Opt.,* vol. 53, no. 34, pp. 8060–8070, 2014.

[55] M. A. Maleki, "Approximate message passing algorithms for compressed sensing," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2010.

[56] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way,* 3rd ed. New York: Academic Press, 2008.

[57] R. Marcia and R. M. Willett, "Compressive coded aperture video reconstruction," in *Proc. 2008 16th European Signal Processing Conf.*, Lausanne, Switzerland, pp. 1–5.

[58] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar, "Compressive light field photography using overcomplete dictionaries and optimized projections," *ACM Trans. Graph.,* vol. 32, no. 4, p. 46, 2013.

[59] A. Mousavi, A. Patel, and R. G. Baraniuk, "A deep learning approach to structured signal recovery," in *Proc. 53rd Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, 2015.

[60] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,". *Appl. Comput. Harmon. Anal.,* vol. 26, no. 3, pp. 301–321, Aug. 2009.

[61] D. Needell and R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Found. Comp. Math.,* vol. 9, no. 3, pp. 317–334, 2009.

[62] J. Y. Park and M. B. Wakin, "Multiscale algorithm for reconstructing videos from streaming compressive measurements," *J. Electronic Imaging,* vol. 22, no. 2, p. 021001, 2013.

[63] R. Saab, R. Wang, and O. Yilmaz, "From compressed sensing to compressed bit-streams: Practical encoders, tractable decoders," *arXiv preprint arXiv*:1604.00700, 2016.

[64] R. Raskar, A. Agrawal, and J. Tumblin, "Coded exposure photography: Motion deblurring using fluttered shutter," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 795–804, 2006.

[65] D. Reddy, A. Veeraraghavan, and R. Chellappa, "P2C2: Programmable pixel compressive camera for high speed imaging," in *Proc. 2011 IEEE Conf. Computer Vision and Pattern Recognition*, Colorado Springs, CO, pp. 329–336.

[66] A. C. Sankaranarayanan, P. K. Turaga, R. Chellappa, and R. G. Baraniuk, "Compressive acquisition of linear dynamical systems," *SIAM J. Imag. Sci.,* vol 6, no. 4, pp. 2109–2133, 2013.

[67] A. C. Sankaranarayanan, L. Xu, C. Studer, Y. Li, K. F. Kelly, and R. G. Baraniuk, "Video compressive sensing for spatial multiplexing cameras using motion-flow models," *SIAM J. Imag. Sci.,* vol. 8, no. 3, pp. 1489–1518, 2015.

[68] A. Secker and D. Taubman, "Lifting-based invertible motion adaptive transform (LIMAT) framework for highly scalable video compression," *IEEE Trans. Image Process.,* vol. 12, no. 12, pp. 1530–1542, 2003.

[69] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. Lensch, "Dual photography," *ACM Trans. Graph.,* vol. 24, no. 3, pp. 745–755, 2005.

[70] P. Sen and S. Darabi, "Compressive dual photography," *Comput. Graphics Forum,* vol. 28, no. 2, pp. 609–618, 2009.

[71] S. Tambe, A. Veeraraghavan, and A. Agrawal, "Towards motion aware light field video for dynamic scenes," in *Proc. 2013 Int. Conf. Computer Vision*, Sydney, Australia, pp. 1009–1016.

[72] J. Tropp et al., "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans.* Inf. Theory, vol. 50, no. 10, pp. 2231–2242, 2004.

[73] N. Vaswani, "Kalman filtered compressed sensing," in *Proc. 15th IEEE Int. Conf. Image Processing*, San Diego, CA, 2008, pp. 893–896.

[74] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4595–4607, 2010.

[75] A. Veeraraghavan, D. Reddy, and R. Raskar, "Coded strobing photography: Compressive sensing of high speed periodic events," *IEEE Trans. Pattern Anal. Mach. Intell.,* 33(4):671–686, Apr. 2011.

[76] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.,* vol. 47, no. 10, pp. 44–51, 2008.

[77] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R. G. Baraniuk, "Compressive imaging for video representation and coding," in *Proc. Picture Coding Symposium*, Beijing, China, 2006.

[78] J. Wang, M. Gupta, and A. C. Sankaranarayanan, "LiSens—a scalable architecture for video compressive sensing," in *Proc. 2015 IEEE Int. Conf. Computational Photography*, Houston, TX, pp. 1–9.

[79] A. E. Waters, A. C. Sankaranarayanan, and R. G. Baraniuk, "SpaRCS: Recovering low-rank and sparse matrices from compressive measurements," in *Proc. Neural Information Processing Systems*, Granada, Spain, 2011, pp. 1089–1097.

[80] J. Wright, A. Ganesh, K. Min, and Y. Ma, "Compressive principal component pursuit," *Information and Inference,* vol. 2, no. 1, pp. 32–68, 2013.

[81] S. J. Wright, R. D. Nowak, and M. A. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.,* vol. 57, no. 7, pp. 2479–2493, 2009.

[82] L. Ying, L. Demanet, and E. Candes, "3D discrete curvelet transform," *Proc. SPIE*, vol. 5914, p. 591413, 2005.

SP

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

Ruhi Sarikaya

The Technology Behind Personal Digital Assistants

An overview of the system architecture and key components

BACKGROUND IMAGE ©ISTOCKPHOTO.COM/ TRAFFIC_ANALYZER; PDA IMAGE ©ISTOCKPHOTO.COM/HUDIEMM

We have long envisioned that one day computers will understand natural language and anticipate what we need, when and where we need it, and proactively complete tasks on our behalf. As computers get smaller and more pervasive, how humans interact with them is becoming a crucial issue. Despite numerous attempts over the past 30 years to make language understanding (LU) an effective and robust natural user interface for computer interaction, success has been limited and scoped to applications that were not particularly central to everyday use. However, speech recognition and machine learning have continued to be refined, and structured data served by applications and content providers has emerged. These advances, along with increased computational power, have broadened the application of natural LU to a wide spectrum of everyday tasks that are central to a user's productivity. We believe that as computers become smaller and more ubiquitous [e.g., wearables and Internet of Things (IoT)], and the number of applications increases, both system-initiated and user-initiated task completion across various applications and web services will become indispensable for personal life management and work productivity. In this article, we give an overview of personal digital assistants (PDAs); describe the system architecture, key components, and technology behind them; and discuss their future potential to fully redefine human–computer interaction.

Digital Object Identifier 10.1109/MSP.2016.2617341
Date of publication: 11 January 2017

Introduction

We are living in the mobile Internet computing cycle. During the past decade, mobile devices have experienced unprecedented growth. According to Statista [65], there are currently more than 4.6 billion mobile phone users in the world, and the number is expected to grow even more moving forward. With this phenomenal increase in volume came technical sophistication and improved capabilities of mobile devices (both on the hardware and software sides), particularly around applications and web services where users can complete a wide array of tasks. As the need and expectation to do more grew, despite improvements, a limited natural user interface has remained as one of the major bottlenecks in interacting with these devices. PDAs (also known as *virtual assistants*) precisely target this problem and have the promise of enhancing a user's productivity by either proactively providing the information the user needs in the right context (i.e., time and place) or reactively answering a user's questions and completing tasks through natural language. Tasks can be related to device functionality, applications, or web services.

Research on PDA technology, however, started much earlier than the emergence of mobile devices. Over the last 20 years, researchers have investigated personalized virtual assistant agents targeting specific domains, including tourism, elder care, device control, and home and office applications [1]–[5]. However, attempts at bringing them to market earlier have failed because of their limited utility.

PDAs have the promise of enhancing a user's productivity by either proactively providing the information the user needs in the right context or reactively answering a user's questions and completing tasks through natural language.

Over the past five years, there has been tremendous investment in PDA technology by both small and big technology companies. Siri [17], [66], Google Now [67], Cortana [68], and Alexa [69] are the major personal assistants in the market today, and they provide proactive and/or reactive assistance to the user. *Proactive assistance* refers to the agent taking an action to assist the user without the user's explicit request. *Reactive assistance* refers to the agent responding to the user's voice or typed command to assist him or her. The number of smartphone users using PDAs increased from 30% in 2013 to 65% in 2015 [70], indicating increased adoption.

PDAs have become a key capability in most smartphones. They are now also deployed in tablets, laptops, desktop PCs, and headless devices (e.g., Amazon Echo), and some are also even integrated into operating systems. These agents are designed to be personal; they know their user's profile, whereabouts, schedules, and so forth. They can proactively start interactions with their user through notifications and system-initiated questions or reactively respond to user requests. User–PDA interactions typically take place via natural language, where the user speaks to the agent as if he or she were speaking to a real human assistant.

PDAs

What is a PDA?

A PDA is a metalayer of intelligence that sits on top of other services and applications and performs actions using these services and applications to fulfill the user's intent. A user's intent could be explicit, where the user commands the system to perform an action, or it could be inferred, where the agent notifies or makes suggestions upon evaluation of one or more triggering conditions it has been tracking. PDAs make use of some core set of technologies, such as machine learning, speech recognition, LU, question answering (QA), dialog management (DM), language generation (LG), text-to-speech (TTS) synthesis, data mining, analytics, inference, and personalization.

Why do we need PDAs?

PDAs are built to help the user get things done (e.g., setting up an alarm/reminder/meeting, taking notes, creating lists) and provide easy access to personal/external structured data, web services, and applications (e.g., finding the user's documents, locating a place, making reservations, playing music). They also assist the user in his or her daily schedule and routine by serving notifications and alerts based on contextual information, such as time, user location, and feeds/information produced by various web services, given the user's interests (e.g., commute alerts to/from work, meeting reminders, concert suggestions). Collectively, these functionalities are expected to make the user more productive in managing his or her work and personal life.

For example, airline travel is a commonly supported scenario by most PDAs. If the user has booked a flight and received a confirmation e-mail along with an itinerary, the PDA scans the e-mail, extracts the flight information, and stores it on the service. On the day of travel, the PDA computes the user's current location using the global positioning system (GPS) on the device, checks the traffic conditions to the airport, and tells the user when to leave for the airport. It also checks the flight status and updates the user if there is a delay, using a flight card as shown in Figure 1. Additionally, it provides weather forecasts for the destination as well as currency conversion rates. Typically, a user has to use multiple applications to go through each of these steps to find out the needed information that is listed on the cards for the travel. None of these atomic steps is significant in isolation, but stitching them together can potentially mark a breakthrough in usefulness to the user. This is the key promise of PDAs.

What is personal about PDAs?

PDAs are expected to be personal. Ideally, the PDA is expected to know who its user is, what its user does, its user's interests, what its user needs, and when and where its user needs it. Despite numerous efforts over the past 20 years to make human–computer interaction personal (particularly around web search), personalization has remained largely broken until recently, not only for PDAs but also for general human–computer interaction due to the four main gaps.

1) *Data*: The system had a limited amount of data to properly model the user and his or her interests. It understands the user based on the experience it provides and the feedback loop it uses.
2) *Computing*: The limited computing power and machine learning were not adequate for modeling the complexity of user behavior.
3) *Interest*: There has been conflict of interest between the user, the platform, and those who pay for the user's attention. The system has not been necessarily prioritizing the user's interests over these other actors.
4) *Content/action*: The system does not support the actions the user wants to perform or does not have the content to serve the user's interests.

During the past seven years, two primary changes have occurred that allowed PDAs to be personal: 1) the increased number of device sensors on mobile phones and 2) the quality and quantity of the user data and digital artifacts (on the device and/or in the cloud) coming from the web services and applications the user accesses. As a result of these changes, it is now possible to represent a user along four axes:

1) *user profile*: user's name, age, gender, parental status, profession, employer, home, work, people in his or her close circle (people graph), favorite places, files, documents, music, photos, and interests explicitly provided by the user
2) *digital activity*: digital artifacts (e.g., calendar, e-mail, social media activity, web searches) on applications and web services
3) *space*: physical location of the user
4) *time*: time at which a specific digital or physical activity takes place.

These four dimensions, when considered together, blend the physical with the digital world and open up new possibilities for powerful inferences and deep user understanding. Inevitably, managing and protecting privacy and security of the user data and information is a major concern, and what has been done in that space is critical, but it is outside the scope of this article.

Mobile device sensors

The computational power and capabilities of mobile phones are increasing every year. The number of built-in sensors on smartphones (e.g., Samsung Galaxy) more than tripled during the past five years [71]. Smartphone sensors measure motion/orientation, GPS coordinates, and many other user and environmental conditions. For example, a device's gravity sensor provides data to infer complex user gestures and motions, such as shake, swing, or rotation. The rich high-precision data coming out of these sensors are made available through application programming interfaces (APIs) and are used in numerous applications and scenarios. The information is sensitive, as it is personal and contextual. Making it available opens up new research areas like fitness and health applications or

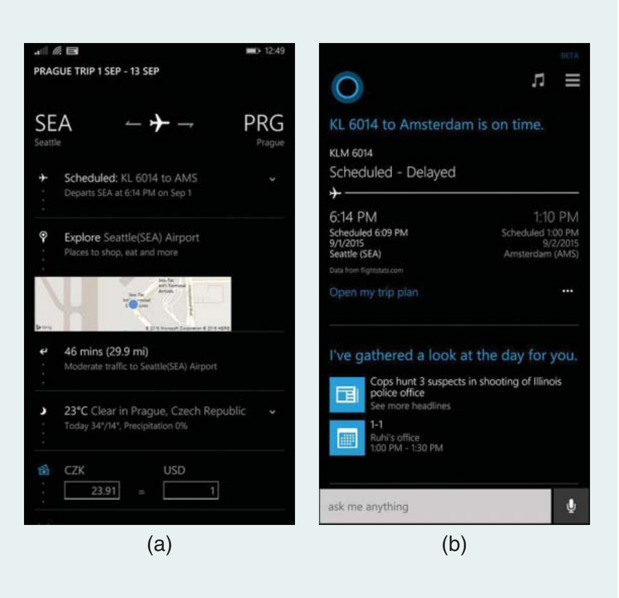


FIGURE 1. The proactive flight cards. (a) Summary and suggestions for the trip. (b) Flight details for the first leg.

opens up new solutions to already existing problems [6]. User experiences that currently exist can also be enhanced by the available data. For example, using activity detection, the PDA can hold the incoming call or send a short message service (SMS) text message to the caller if the user is biking or it can turn up the volume if the user is climbing stairs/walking.

The three main mobile platforms (Android, iPhone operating system, Windows) support four broad categories of sensors on mobile devices.

1) *Motion*: This sensor set includes accelerometers, gravity sensors, gyroscopes, and rotational vector sensors. They measure acceleration and rotational forces along three axes. They measure movement and orientation of the device.
2) *Environmental*: These sensors measure various environmental conditions, such as ambient air temperature and pressure, illumination, and humidity. This category includes barometers, photometers, magnetometers, and thermometers.
3) *Position and location*: These sensors measure the physical position and location of a device. This category includes orientation sensors and magnetometers. The magnetometer can determine the rotation of the device relative to magnetic north. It can also detect magnetic fields around the device. GPS and Wi-Fi (not really a sensor in the traditional sense) determine the location of the device.
4) *Proximity*: This sensor detects whether the phone is brought near the face during a phone call. This functionality disables the touch screen, preventing inadvertent input to the phone from the user's face and can also save battery power.

System architecture

The scenarios that the PDAs support can be divided into two main categories: 1) proactive and 2) reactive assistance. The conceptual agent architecture designed to support these two modes of assistance is shown in Figure 2. The system

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

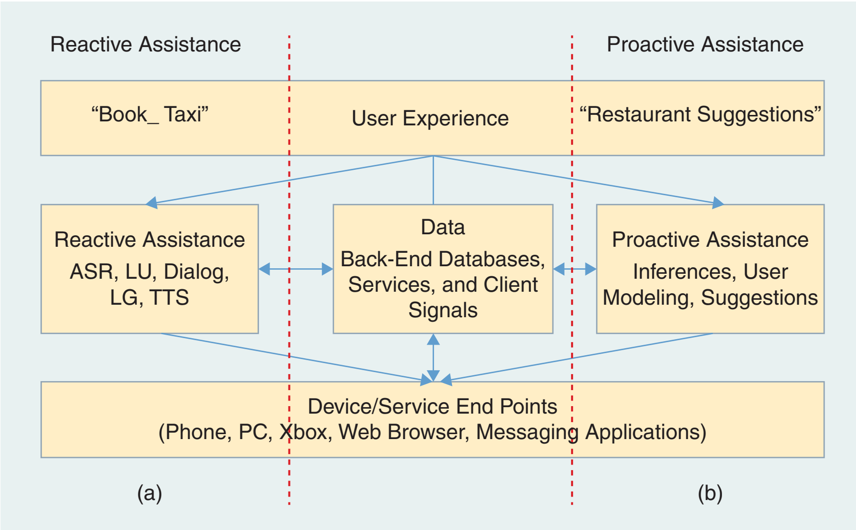


FIGURE 2. The personal digital agent architecture for (a) reactive assistance and (b) anticipatory computing.

architecture depicts proactive and reactive user experiences, data, and service end points. Reactive assistance is shown in Figure 2(a), where the user issues an explicit natural language command (e.g., "book me a taxi") to the agent. The user request is handled through a set of reactive assistance components, such as speech recognition, LU, and DM. The data coming from various back ends, and applications are served to the user according to the constraints specified in the natural language query. The experience (reactive and/or proactive) can be served in one or more of the different device or service end points.

Proactive assistance [Figure 2(b)] involves anticipatory computing, where the personal digital agent does things in a contextual manner (i.e., at the right time and place) that it expects is valuable to the user without an explicit user request. Proactive assistance makes use of inference, user modeling, and ranking to power experiences. Back-end data, device, applications, and web services signals are leveraged for proactive inference and triggering.

Even though proactive and reactive parts of the current PDA architectures are built in isolation, in principle they can use a single architecture to enable both types of experiences.

Even though proactive and reactive parts of the current PDA architectures are built in isolation, in principle they can use a single architecture to enable both types of experiences. In fact, most proactive scenarios have reactive extensions and vice versa. For example, if the user makes a restaurant reservation (reactively), the agent may (proactively) suggest a movie after the dinner or may offer to book a cab to take the user to the restaurant. Data and context are shared between the two assistance modes. Next, we focus on the proactive system architecture and the components that power proactive scenarios.

Proactive assistance

Proactive assistance is based on the theory of proactivity that describes user desires and a model of helpfulness [7]. The goal is to provide assistance to automate tasks or further the user's interests for things he or she cares about, all within context, without explicit user request [8]. To achieve that, the agent is designed to possess a set of attributes; it should be valuable in that it advances the user's interests and tasks, while not interfering with the user's own activities or attention unless it has the user's explicit approval. It should be unimposing. The agent should be transparent in what it knows about the user. It should be anticipatory and know the future needs of the user and bring opportunities to the surface. The agent should also continuously learn and refine its decisions from the feedback signals it receives regarding the actions it takes. These principles put the user at the center, and the agent's actions are considered valuable only if they ultimately add value for the user. Proactive assistance operates on the proactivity continuum [31], which ranges from zero to full automation, allowing for the following scenarios:

- do it yourself (no help from the agent)
- user tells the agent what to pay attention to (notifications/alerts)
- agent infers user's habits/patterns and makes suggestions (inference/suggestions)
- agent makes decisions and takes actions (full autonomy on task decisions/executions).

Most of the currently supported proactive scenarios are notifications/alerts and suggestions. Even though there is some preliminary work, none of the agents in production supports autonomous decision making and action taking on behalf of the user without confirmation.

The proactive agent system architecture is shown in Figure 3. Signals coming from web services, device sensors, and the user's profile are processed, where processing includes parsing, enriching, and filtering to merge device and service data. The next step is aggregation, which joins the processed data streams through time and space (i.e., location) about the user's whereabouts and actions/tasks done at specific times and places. This step blends the physical and digital worlds and allows for powerful inferences that capture repetitive behavior and events in both worlds. The signals are used to make inferences and train machine-learned models for modeling the user and his or her interests. The same set of signals is also used to set rules for notifications and alerts the user wants the agent to serve. The models and rule recipes are deployed to a run time environment. Once proactive scenarios are deployed in production, capturing and feeding back user behavior signals regarding notifications, alerts, and

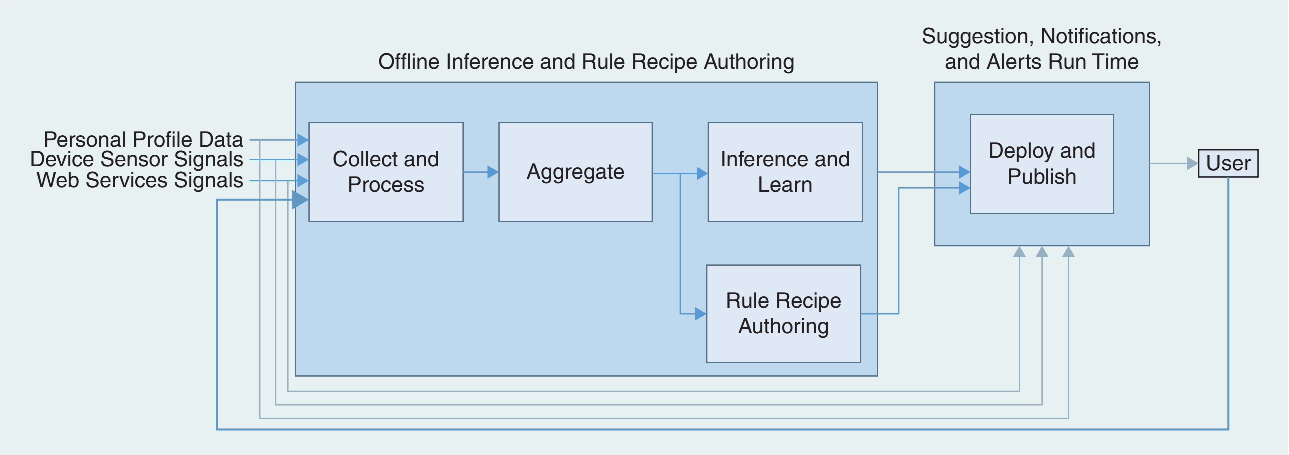


FIGURE 3. The proactive assistance system architecture.

suggestions are essential for the proactive agent to learn and adapt to the user.

Notifications and alerts

In the notifications/alerts category, the system allows the user to set rules to define the triggers for certain actions. If the triggering condition evaluates to TRUE, the action is executed. The rules are defined over a set of signals. These signals are produced by an information channel that can be evaluated by the proactive agent. These channels represent many types of information, such as date/time and location, as well as constantly updated data feeds generated by various web services, which include weather, sports, news, finance, and entertainment. For example, one can create rules to obtain an alert when the Seattle Seahawks score a touchdown. A user can set a rule to be reminded of his or her mother's birthday. It is also possible to combine these signals to formulate more complex triggering rules. For example, a specific flight departure time, a user's physical location, and a commute time to the airport can all be used to trigger an alert that reminds the user that it is time to leave for the airport. Once the trigger rule is set, the agent monitors the signals from the corresponding information channels to evaluate the rule. If the rule evaluates to TRUE, the agent takes an action. The actions are communicated to the user in a target device-specific manner, which could be a proactive entity card, SMS, or even a phone call. This type of proactive agent programming falls under the if-then recipes [60], in which simple rules allow users to control many aspects of their digital life.

Inference and suggestions

In the suggestions category, the agent infers the user's habits and routines by reasoning over his or her past behavior and makes a personalized recommendation to the user with the goal of furthering the user's interest. For example, knowing that the user watched comedy movies featuring a particular actor in the past, the agent may suggest a new comedy movie featuring that same actor in the future. The agent can also suggest new experiences based on the logical sequencing of different yet related (through time or location) events. For instance, if the user made a restaurant reservation in a metropolitan city downtown, the agent may suggest nearby parking places. Through inference, the agent can learn certain facts about the user, by reasoning over the user's whereabouts and movement patterns through time and location. For example, the user's home and work location could be inferred by joining GPS data with time over several weeks. If the user is spending most or all of his or her time between 9 a.m. and 5 p.m. during weekdays at a specific location over several weeks, that is likely to be the user's work location. Likewise, the user's commute hours between home and work could also be inferred from combining home and work location with the GPS data during the likely morning and evening commute hours over several weeks. This inference is used to proactively show the traffic commute cards around the time the user typically commutes to/from work (or home).

The key questions here are determining the type of suggestion and when to do it, because there is an associated cost with the suggestion (if the action, relevance, or timing is wrong). To get around the cold start problem (if the agent does not have access to the user's past activity through a feedback loop or the user is accessing the PDA for the first time), the user is also given the ability to teach the agent his or her interests from a precompiled list of topics, including news, sports, finance, technology, dining, and entertainment. The decision for taking a proactive action is driven by a machine-learned model, given the costs and benefits as constraints. The machine-learned model combines a set of information in the user's profile, demographic and content-based profiles, and online user behavior signals (such as click through, dwell time, and dismissal), along with the user's recent relevant activity (e.g., similar content searches), which are captured in the history variable h in (1). This is used to model whether a specific user u will like the specific suggested entity e. Standard machine-learning techniques, such as maximum entropy models [35], gradient boosted decision trees [55], and deep learning techniques [48], are used to incorporate both user-specific online and offline signals to estimate the probability that the user is expected to like the suggested entity

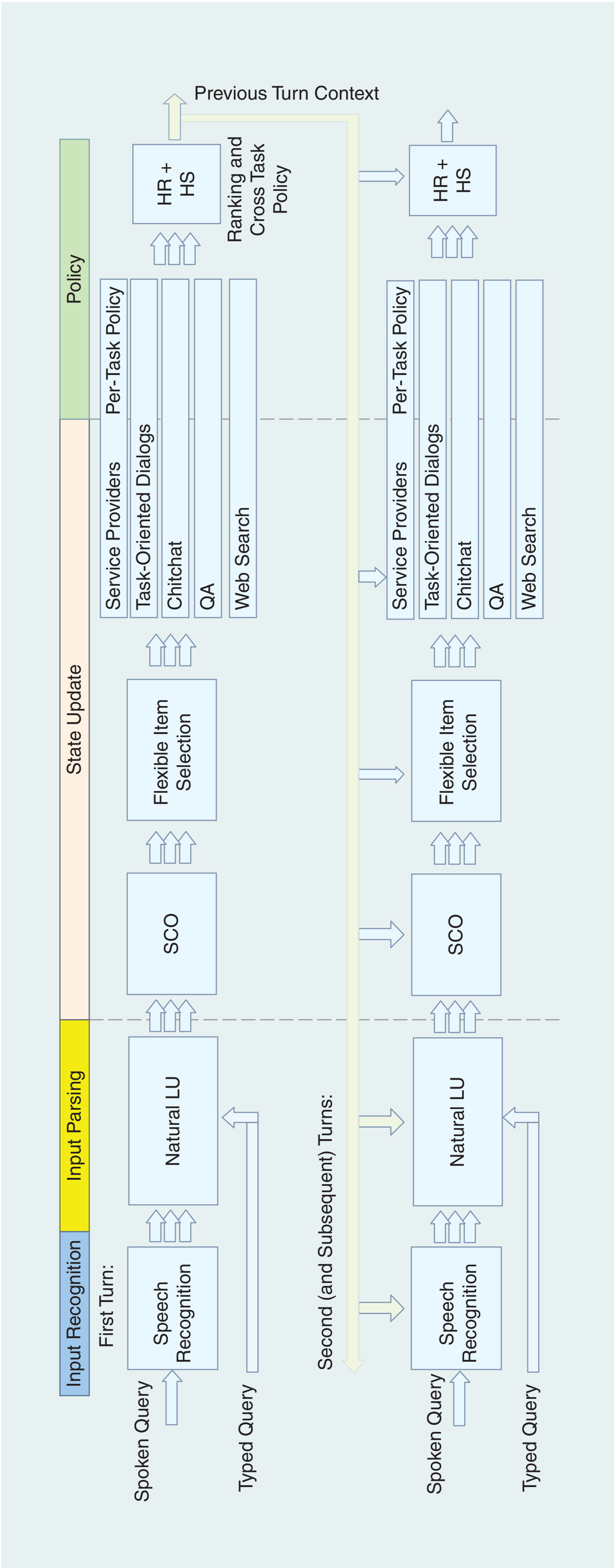


FIGURE 4. The reactive assistance architecture: conversational understanding system architecture.

(i.e., content). Within the maximum entropy modeling framework, the probability is computed using

$$P(o|\, e, u, h) = \frac{e^{\sum_i \lambda_i f_i(o,e,u,h)}}{\sum_{o'} e^{\sum_{j \in J} \lambda_j f_j(o',e,u,h)}}. \tag{1}$$

Here, o denotes outcome (e.g., like or dislike). Notice that the denominator includes a sum over all possible outcomes, o', which is essentially a normalization factor for probabilities to sum to 1. The functions f_i are usually referred to as *feature functions*, or simply *features*. These binary feature functions are given as

$$f_i(o,e,u,h) = \begin{cases} 1, & \text{if } o = o_i \text{ and } q_i(e,u,h) = 1 \\ 0, & \text{otherwise} \end{cases}, \tag{2}$$

where o_i is the outcome associated with feature f_i and $q_i(e,u,h)$ is an indicator function on the user, suggestion, and history. The model parameters λ_i are learned on labeled data, which capture the user's response (e.g., like or dislike) for the suggested content in the past.

The specific features of the suggested entity (e) include the estimated value of suggestion type, value of suggestion instance, timing of the suggestion, cost of mistake, cost of interruption, and urgency (time sensitivity) of the suggestion. Learned thresholds on the model outputs govern the number of suggestions of each type that can be displayed concurrently, the maximum frequency for suggestions of each type, and the permitted or prohibited modalities of each suggestion type. The thresholds for acting, asking, suggesting, or doing nothing are established from a range of default values according to user-stated advice and elicited initial preferences from the configuration wizard.

Reactive assistance

Reactive assistance is traditionally known as the *conversational understanding system*. The conversational understanding system for PDAs spans a wide spectrum of domains, including goal-/task-oriented dialogs [53], [16], chitchat, QA [37], and classical web search answers. The conversational understanding system also handles additional input modalities besides speech, such as typing and/or touch. Each of these domains is commonly referred to as an *answer* (*A*). Some of the domains involve device functionality (e.g., alarm, SMS, calling, note, reminder), while others may involve web services and applications (e.g., directions to a particular location, movie hours at a theater, factoids, stock prices, weather).

The reactive assistance system architecture is shown in Figure 4. The user submits a request to the PDA to perform a task or seek information using one of the modalities, and the agent interprets the request and generates a response. For voice queries, the first step is to recognize the spoken words [9], [10]. The LU component takes the speech transcription (or the text input if the user types) and performs a semantic analysis to determine the underlying user intent [11]–[14], [17]. The user's intent could be related to information search, QA, chitchat, or task-oriented specialist dialogs. Because PDAs support multidomain and multiturn interactions, multiple alternate semantic analyses (typically at least one for each domain) are generated in parallel for late binding on the user's intent [15]. These semantic analyses are sent to the dialog state update component, which includes slot carryover (SCO) [38], flexible item selection from a list [39], knowledge fetch from the service providers, and dialog hypothesis generation [15], [16]. Note that in this framework, we consider chitchat, QA, and web search as an additional set of LU domains. All the dialog hypotheses are ranked by the hypothesis ranking (HR) module. The top hypothesis is selected by the hypothesis selection (HS) module by taking the provider responses (i.e., knowledge results) into account [18]. The top dialog hypothesis (along with the ranked dialog hypothesis distribution) is the input to the dialog policy component, which determines the system response based on the scenario and business logic constraints. Typically, for voice input, the agent speaks the natural language response via the TTS synthesis engine [19].

The reactive assistance behavior is governed by (3). The goal of the reactive agent is to provide the best system response $\hat{R}$ to a given user query, Q. The system response, R, consists of a dialog act, which includes system action (e.g., information to be displayed, question to be asked, or action to be executed), natural language prompt, and a card in which the response is displayed

$$\hat{R} = \underset{R}{\operatorname{argmax}} \{P(R \mid Q, B_{A_1}, \ldots, B_{A_N}, \bar{B})\}, \tag{3}$$

where B_{A_1} denotes the current belief about the dialog state of the answer A_1 (e.g., weather, alarm, places, reminder, sports, etc.) after processing query Q, and $\bar{B}$ shows the system's belief about the state of the interaction across all answers for the current session. In practice, it is hard to solve (3). Instead, a suboptimal solution can be achieved with the assumption that, given the query Q and the beliefs for the dialog states of the individual answers A_1 through A_N, the per answer response is conditionally independent

$$\hat{R} = \underset{R \in R_1, \ldots, R_N}{\operatorname{argmax}} \{P(R \mid Q, B_{A_1}, \bar{B}), \ldots, P(R \mid Q, B_{A_N}, \bar{B})\}. \tag{4}$$

Here, $P(R \mid Q, B_{A_i}, \bar{B})$ denotes the probability the system assigns to response (R) generated by answer A_i, given the answer's belief about its dialog state and the system's belief, $\bar{B}$. This formulation allows the individual answers to manage their own dialog state and generate their own responses in parallel. Therefore, it is possible to scale to many domains and answers without substantially increasing the overall system response latency. The HR component operates as a metalayer, arbitrating between different answer responses, given its belief (i.e., $\bar{B}$) about the state of the overall interaction [15]. Next, we will briefly describe the key components in Figure 4.

Speech recognition

The speech recognition component maps the human speech represented in acoustic signals to a sequence of words represented in text. Let X denote the acoustic observations in the form of feature vector sequence and Q be the corresponding word sequence (i.e., query). The speech recognition decoder chooses the word sequence, $\hat{Q}$, with the maximum a posteriori probability according to the fundamental equation of speech recognition [9]:

$$\hat{Q} = \underset{Q}{\operatorname{argmax}}\, P(X \mid Q) P(Q), \tag{5}$$

where $P(X \mid Q)$ and $P(Q)$ are the probabilities generated by the acoustic and language models, respectively. Traditionally, speech recognition systems are trained to optimize the lexical form. However, displaying the grammatically and semantically correct version of the output (i.e., display form) has become an important requirement for PDAs, because it makes it easy for the user to infer whether the system correctly heard and recognized the spoken query. For example, the following two speech recognition outputs are lexically equivalent:

- how is the traffic on u s one oh one (lexical form)
- how is the traffic on US 101 (display form).

However, proper tokenization in the second hypothesis provides a valuable hint that the agent understood what the user meant. Typically, the tokenization is applied as a separate postprocessing module after running the speech recognition decoder.

In recent years, advances in deep learning and its application to speech recognition have dramatically improved state-of-the-art speech recognition accuracy [9], [10], [20], [21]. Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. These advances played a key role in the adoption of PDAs by a large number of users making it a mainstream product.

LU

The problem of LU for PDAs is a multidomain, multiturn, contextual query understanding [17], [22]–[25], subject to the constraints of the back-end data sources and the applications in terms of the filters they support and actions they execute. These constraints are represented in a schema. In practice, while LU semantically parses and analyzes the query, it does not do so according to a natural LU theory [26]; rather, parsing and analysis are done according to the specific user experience and scenarios to be supported. This is where the semantic schema comes into play, as it captures the constraints of the back-end knowledge sources and service APIs, while allowing free form of natural language expression to represent different user intents in an unambiguous manner.

There are two main approaches to LU: rule based and machine learned [11], [32]. The rule-based approach is about hand authoring a set of rules to semantically parse the query [27]. It can also be used for addressing the errors and disfluencies introduced by a speech recognizer [28], [49]. State-of-the-art systems use machine-learned models for LU [12], [17], [23], [24]. In a commonly used LU architecture, a query is first classified into one of the supported domains (or a catch-all domain such as web). Typically, support vector machines, boosted decision trees, maximum entropy models [35], or neural networks [13], [33], [34] are used for modeling. These classifiers are trained in either binary or multiclass mode depending on the design choices [12], [23], [24]. Under each domain, there are a domain-specific intent and slot model. Intents and slots can be shared across different domains. A domain can be considered as a collection of related intents, which do not have any conflict. For example, a weather domain contains check_weather and get_weather_stats intents. Intent detection uses the same machine-learning techniques listed previously, and it is framed as a multiclass classification problem. Slot tagging is considered as a sequence classification problem. Conditional random fields, maximum entropy Markov models [35], and, more recently, deep learning techniques [12], [24] are used for slot tagging. LU models are trained in a supervised fashion using labeled data and properly weighted lexicons [36]. When a user issues a query, domain and intent classifiers are run to determine the domain and intent of the query, and the slot tagger tags semantic slots. Tagged slots are resolved into canonical values, and in some cases multiple slots are combined into a parameter. Parameters are used either to fetch results from the knowledge back end or to invoke an application API.

The goal of the LU component (for answer A_i) is to convert each input query into a set of semantic frames, F_{A_i}, given the context represented in the current beliefs about the dialog state B_{A_i}, that is, we seek

$$\hat{F}_{A_i} = \underset{F}{\operatorname{argmax}}\{P(F|\hat{Q},B_{A_i})\}. \tag{6}$$

Semantic frame, F_{A_i}, for answer A_i encapsulates the semantic meaning of a query with a tuple of domain, intent, and slot list:

$$\langle \text{DOMAIN, INTENT, SLOTS} \rangle.$$

The slots are a list of key-value pairs:

$$\langle \text{SLOT\_NAME, SLOT\_VALUE} \rangle.$$

The data structure for semantic frame, shown in Figure 5, is used to combine the different pieces of semantic analysis

```
<Domain score="0.8956">reminder</Domain>
    <Intent score="0.7949">create_single_reminder</intent>
  <Slots>
     <reminder_text ="call my mom" />
     <start_time <RawValue="9 am">
            <PropertyGroup Name="timex3">
               <Property Name="value" Value="am" />
               <Property Name="comment" Value="am" />
            </PropertyGroup>
      </start_time>
  </Slots>
```

FIGURE 5. The semantic frame for the query: "remind me to call my mom at 9 a.m."

generated by the domain, intent, and slot models to represent the complete semantic understanding of the query.

Some of the key PDA experiences that are handled are multiturn in nature. Without using contextual information, the queries could be ambiguous and potentially interpreted differently. For example:

- (turn 1) how is the weather in New York (weather)
- (turn 2) what about the weekend (weather).

Here is another scenario, where we observe the exact same query in the second turn:

- (turn 1) how is my schedule (calendar)
- (turn 2) what about the weekend (calendar).

Interpreting the follow-up queries in isolation is difficult, as they are ambiguous and require context for proper interpretation. LU models are built in a contextual manner to solve this problem [15], [22]–[24]. To handle multiturn interactions, in addition to basic domain, intent, and slot models, one needs to build context carryover models to help with state tracking [38] and on-screen selection models [39], as shown in Figure 4, for selecting items from a list of results presented to the user in follow-up turns.

QA

Because users are expecting PDAs to answer any question, open-domain QA [56] is another scenario that is handled by all PDAs to differing degrees. Examples of open-domain factoid QA include the following questions:

- How old is Bill Gates?
- How tall is Mount Everest?
- Who directed *Avatar*?

The answers to these questions are precise short phrases.

- Bill Gates is 60 years old.
- Mount Everest is 8,848 meters high.
- James Cameron directed the movie *Avatar*.

QA has had a long history [29] and has seen rapid advancement in the past decade, spurred by government-funded programs that required system building, experimentation, and evaluation of systems [30]. Advancements in search engine technology, such as query formulation and query-document analysis through click logs, have also contributed to innovation in QA [52].

The system architecture for a typical QA system is shown in Figure 6. For a given query directed at a PDA, the QA system first classifies the query into one of the question types (typically ten to 15). Note that the query may not be a question. Even if it is a question, it may not be supported by the QA system. These categories are also included as part of the question classification step. The answer-candidate generation step processes the question, generates various alternate formulations of it, and queries the knowledge sources, which include a knowledge graph, web documents, Wikipedia, and search engines. The answer-candidate ranking step extracts a number of features for each question/answer pair and applies a ranking model to rank and assign confidences to each answer candidate [56]. The overarching principles of QA systems are massive parallelism, confidence estimation, and integration of shallow and deep knowledge from many knowledge sources [37].

Knowledge back ends

A significant part of PDA scenarios is about accessing knowledge and entities. For example, when a user asks for a factoid regarding a movie or director, LU models tag such slots (i.e., facets) as movie name, release date, actors, and directors. Slots are used to build a query sent to these knowledge bases to fetch the relevant entity and relationships for which the user is looking. These knowledge bases store factual information in the form of entities and their relationships, covering many domains (people, places, sports, business, etc.) [57]. The entities and their relationships are organized using the World Wide Web Consortium Resource Description Framework (RDF) [40]. An RDF semantic knowledge base (also known as a *semantic knowledge graph*) represents information using triples of the form subject-predicate-object, where in graph form, the predicate is an edge linking an entity (the subject) to its attributes or another related entity, as seen in Figure 7. A popular open-source RDF semantic knowledge base is Freebase [46]. Other RDF semantic knowledge bases that are much larger in size are Facebook's Open Graph, Google's Knowledge Graph, and Microsoft's Satori. Both Google's Knowledge Graph and Microsoft's Satori knowledge graph have over 1 billion entities and many more entity relationships. They power entity-related results that are generated by Google's and Microsoft Bing's search engines. Recently, there has been a surge of interest in exploiting these knowledge sources, especially the RDF semantic knowledge bases, to reduce the manual work required for expanding conversational systems to cover new domains, intents, or slots [41]–[43], [51] or improving existing experiences [44], [45].

DM and policy

Many of the reactive scenarios enabled in PDAs require spoken DM to handle a wide range of tasks and domains. DM is at the bottom of the reactive stack, where all the information from upstream components is consolidated and the final decision about the system response is made and communicated to the user.

Much of the research on spoken dialog systems in academia has targeted single-domain applications [47], [50], where the problem of accurately tracking the user's goal (e.g., finding restaurants that satisfy a number of user constraints) has received considerable attention in the literature [53]. The primary line of research has been the statistical modeling of uncertainties and ambiguities encountered by dialog managers due to speech recognition and LU errors along with ambiguity in natural language expressions. Included among the most successful statistical approaches are graphical models that are concerned with decision making with delayed rewards [53]. However, large-scale production systems such as PDAs pose a different set of problems. The large number of supported domains, integration of task-oriented dialogs, QA, chitchat, and web answers and managing conversation in a coherent way pose new challenges

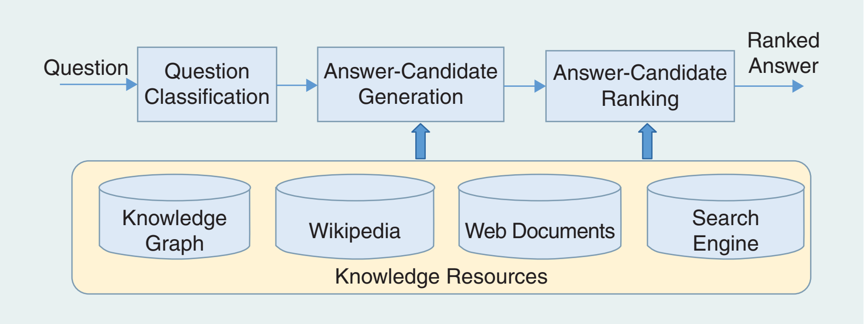


FIGURE 6. The QA system architecture.

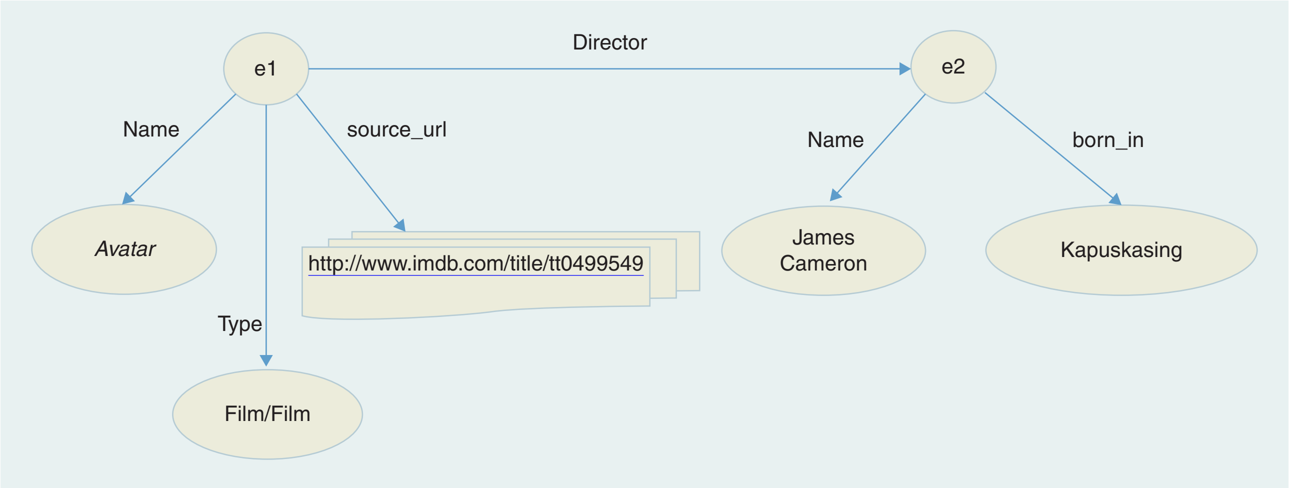


FIGURE 7. An example of part of a semantic knowledge graph representing the relationships, described as RDF triples, between the entities James Cameron (e1) and the film *Avatar* (e2).

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

[25]. Mixing different modalities to complete the tasks is another challenging area for PDAs.

In PDAs, a dialog manager supports execution of a variable number of goal-oriented tasks [16], [54]. Tasks are defined in terms of information to be collected from the user, corresponding LG prompts, and interfaces to resources (such as data hosted in external services and applications) that will execute actions on behalf of the user. As shown in Figure 4, at each turn, the dialog state is updated, taking into consideration the multiple LU results across different turns. SCO [38] does contextual carry over of slots from previous turns, using a combination of rules and machine-learned models with lexical and structural features from the current and previous turn utterances. Flexible item selection uses task-independent, machine-learned models [39] to handle disambiguation turns where the user is asked to select between a number of possible items. The task updater module is responsible for applying both task-independent and task-specific dialog state updates. Task-dependent processing is driven by a set of configuration files, or task forms, with each form encapsulating the definition of one task. Using the task forms, this module initiates new tasks, retrieves information from knowledge sources, and applies data transformations (e.g., canonicalization). Data transformations and knowledge source lookups are performed using resolvers [16], [54]. Dialog policy execution is split into task-specific and global policy. The per-task policy consists of analyzing the state of each task currently in progress and suggesting a dialog act to execute. The dialog acts include show results, disambiguation, prompt for missing value, prompt for no results found, start over, go back, cancel, confirm, complete the task, and so forth, in accordance with the ranked semantic frame output. The output of the task updater module is a set of dialog hypotheses representing alternative states or dialog actions for each task in progress. The dialog hypotheses are ranked using HR [15], [18], which generates a ranked order and score for each hypothesis. This acts as a pseudobelief distribution over the possible dialog/task states. HS policy selects a top hypothesis based on contextual signals, such as the previous turn task, rank order, and scores as well as business logic.

It is generally difficult to empirically evaluate the quality of proactive and reactive user experiences for PDAs.

HR uses an implementation of LambdaMart [55] to rank hypotheses. Previously, various approaches have also been presented for reranking for spoken LU [58] but have focused on single-domain applications.

LG

Once the system response is determined by the dialog manager, it is communicated to the user in a natural way. If the query is a speech query, a spoken system response is returned. If the query is typed, there is no spoken system response but rather a natural language system response that the user sees on the screen as a card. In either case, the natural LG component receives the dialog output in the form of a system response along with the dialog state, which encapsulates the state of the interaction between the user and system (e.g., current turn identification, whether the user has prompted for the same information before) and generates a natural and grammatical utterance to convey the system response. There are a number of factors that feed into the LG design, such as information presentation, presenting enough information (to give a good overview of the state of the task) versus keeping the utterances short and understandable, handling error states, and repeated tries [59].

There are three main approaches to LG: 1) template-based, 2) rule-based (linguistic), and 3) corpus-based approaches [59], [61]. Most of the PDAs use the template-based approach, because comparatively less effort is needed to develop and maintain the templates. The template-based LG module typically starts out from a semantic representation (e.g., semantic frame), generating "QFC in Redmond is open from 7:00 a.m. to 10:00 p.m." in response to the user query, "Is QFC in Redmond open today?" For example, STOREHOURS: [PLACENAME("QFC"), LOCATION("REDMOND")] associates it directly with a template, such as [PLACENAME] in [LOCATION] is open from [TIMEBEGIN] to [TIMEEND], where the gaps represented by [PLACENAME], [LOCATION], [TIMEBEGIN], and [TIMEEND] are filled by looking up the relevant information in the dialog state. The TTS engine consumes the LG output and synthesizes the text into speech [19].

Metrics and measurement for PDAs

It is generally difficult to empirically evaluate the quality of proactive and reactive user experiences for PDAs. PDAs are complex systems with many components in the system stack, spanning client and multiple cloud services, and it is hard to separate any one component from the rest. Each component has its own functional and quality metrics. Metrics could be offline, measured with sampled data sets, or online, measured with actual user traffic of the live system.

Component metrics

The following are the offline component metrics tracked individually to improve the quality of each component. They are computed using a set of ground truths generated by human judgments:

- *speech recognition*: word error rate (WER), sentence error rate, slot WER, keyword WER, semantic WER (for non-search tasks)
- *LU*: domain accuracy, intent accuracy, precision/recall, slot F1 measure, semantic frame accuracy
- *dialog*: dialog state tracking accuracy—in a distribution (e.g., N-best list) of dialog state hypotheses, percent accuracy of the top-ranked hypothesis, selection accuracy, SCO accuracy
- *LG:* mean opinion score (MOS) of the LG quality, bilingual evaluation understudy score
- *knowledge*: knowledge relevance, coverage, precision/recall
- TTS *synthesis*: MOS, intelligibility, expressiveness

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

- *proactive suggestions/notifications*: precision/recall.

In Figure 8, we show LU accuracy for domain, intent, slot tagging (F1 measure), and semantic frame for some of the reactive experiences supported in Cortana for a sample training and test data set [25], [45]. The domain, intent, and slot accuracy is around 90% across domains. The average semantic frame accuracy is 82%. Note that semantic frame combines domain, intent, and slots; therefore, errors in these components contribute to the semantic frame error rate.

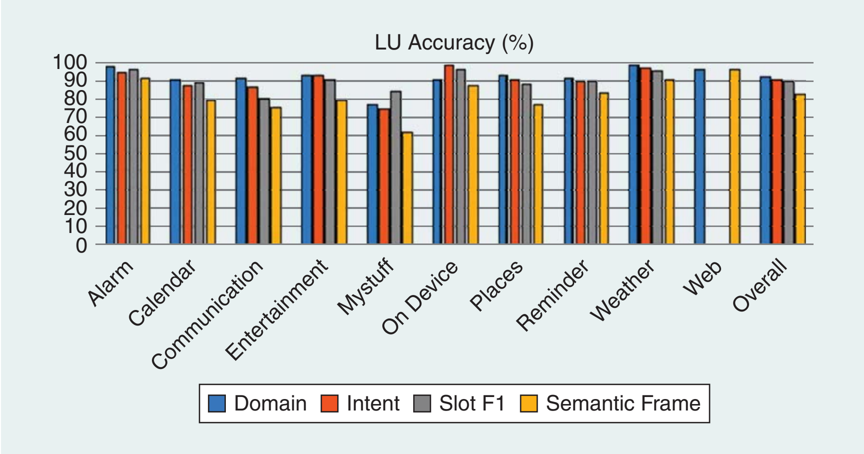


FIGURE 8. The LU domain, intent, slot, and semantic frame accuracy.

In Figure 9, we show the impact of HR in picking the right hypothesis over the LU model confidences. HR improves in picking the correct semantic frame by 2%. HR has the full view of all the LU analyses coming from different domains, and so it can arbitrate between competing hypotheses.

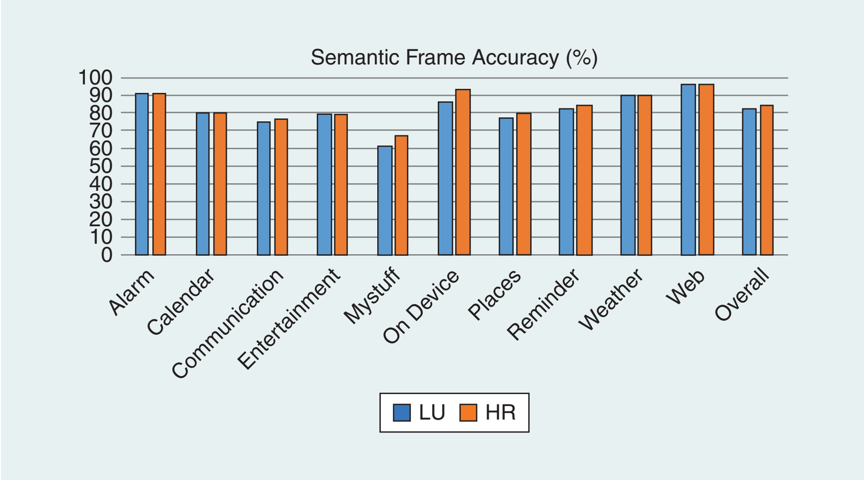


FIGURE 9. The semantic frame accuracy for LU versus HR.

End-to-end quality metrics

The fact that individual component accuracies are high may not mean that the PDA, as a product, has high accuracy. There are several factors contributing to this. For example, speech recognition may not be accurate even if LU is accurate, and knowledge results may not be relevant. There may be operational service reliability issues, back-end availability, network communication issues, robustness of wake-up word detection, and so forth, which all contribute toward quality of the user experience with the product. Moreover, a user's intent could be understood by the LU component, but the underlying application or service may not support that intent. Integrating different client and service components is a challenging software engineering problem, as it uncovers numerous scenario, design, service, and client shortcomings, which take the most time in improving the system. Therefore, end-to-end (E2E) product quality metrics are critical for the success of the product, as they correlate well with the actual user experience. They are also used for evaluating the contribution of the individual components to the overall product experience based on the analysis of how much each component contributes to the E2E error rates. Some of these metrics are as follows.

- *E2E accuracy* is measured through human judgment of query–response pairs on a five-point scale, where the user is shown a screenshot containing the system response, similar to the ones in Figure 10. Human judges assign a score between 1 (terrible) and 5 (perfect) to each

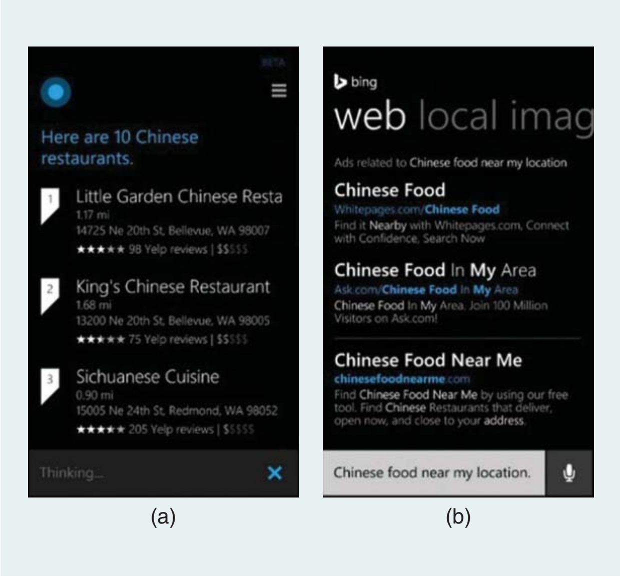


FIGURE 10. The system responses for the query "Chinese food near my location." (a) The correct result from the places domain and (b) the web search result.

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

query response pair. Success includes the ratings of 3 (okay), 4 (pretty good), and 5 (perfect).
- *Side-by-side* (*SBS*) compares system A with system B, where the two systems could be the same system (or competitive systems) at different points in time differing in updates and improvements. SBS is an A/B test on a five- or seven-point scale, where human judges pick one system over the other based on the results shown to the judge. For example, in Figure 10, we show two PDA system responses for the query "Chinese food near my location." In an SBS evaluation, the human judges compare the two responses as to whether the left/right response is better than the other on a scale of $+3$ to -3, where 0 shows that they are equal. SBS is a more sensitive metric compared to E2E accuracy.
- *Online user/system behavior-based metric* measures the user satisfaction and dissatisfaction with the PDA experiences. It is a model that uses a set of feedback signals from the user and the system that correlates with the quality of the user experience. The signals include actions executed, query reformulation, total elapsed time for task completion, landing page dwell time, click through, start over, cancel, click back rates, and latency.

It is quite challenging to evaluate PDAs, as they provide a wide range of experiences, including voice commands, task completion, chitchat, QA, and web search. Therefore, success/failure signals for online measurement could be quite different [62]. An instance where no click has occurred on the screen (i.e., abandonment) in one experience may mean user satisfaction (e.g., showing a correct weather card), but it may mean dissatisfaction in another domain (e.g., showing a list of restaurants and prompting the user to select), where the user leaves satisfied in the former but dissatisfied in the latter. There are also additional metrics used for business and overall product success, including user count, daily/monthly average users, sessions per user, query volume, unique query count, and number of proactive page views.

In Figure 11, we show the E2E query–response pair accuracy. In the figure, E2E Success* denotes the accuracy after leaving out the use cases the scenario is not designed to handle in the first place. For example, the user wants to delete an alarm on the PDA, but the scenario is not supported by design. Instead, the PDA shows either an irrelevant web search result or invokes the alarm application.

In Figure 12, we show the distribution of user dissatisfaction with regard to the sources of error across different components of the system. The numbers are based on feedback of about 10,000 real users. LU along with unsupported system action (e.g., user wants to delete an alarm but system does not support that action) are the biggest sources of user dissatisfaction. Frequent fallback to web search (i.e., text links in search results) when the system does not a have precise answer, a lack of LG (for scenarios where the user expects the system to talk), and speech recognition errors are the main buckets of user dissatisfaction with PDAs.

Technology and user experience challenges

While the functionality and types of tasks a PDA can perform are quite diverse and users find great value in using them, there are still a number of user experience and technical challenges that have yet to be addressed properly. We categorize these challenges into the following groups.

User experience challenges

User experience challenges include the challenges in the following list.
- *Operation errors*: There is a discrepancy between the user's mental model of the PDA's functionality and scenario coverage versus the actual PDA functionality. Users are often unaware of the total extent of the operations a PDA can perform. Users may not understand how to use the PDA application or what they need to say to get the result they desire. Current user interfaces lack the ability to

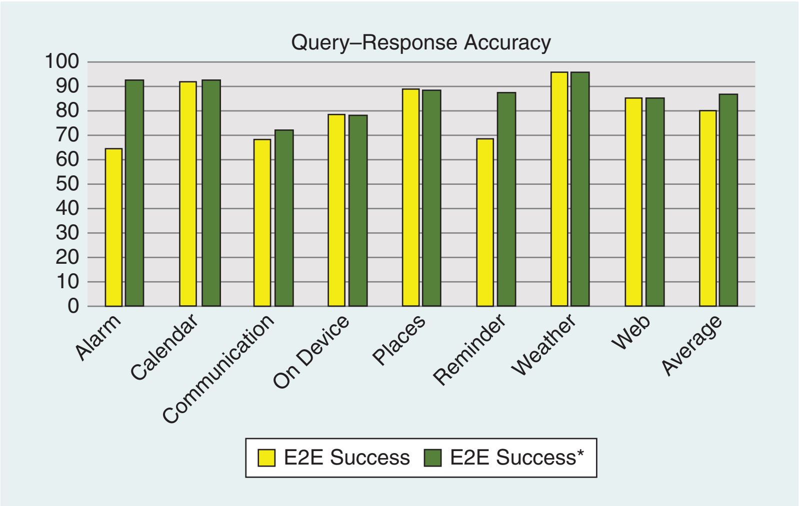


FIGURE 11. The E2E query–response accuracy. E2E Success* denotes the accuracy after leaving out the use cases that the scenario is not designed to handle in the first place.

provide sufficient information about how to use the system and intuitive sequence of operations to complete a task. PDAs are also not able to adapt to different user profiles and their way of operations.

- *Lack of competence*: PDAs are not at a level to reliably decide when to help the user, what to help the user with, and how to help him or her. This also creates a trust issue between the user and the agent, and the user may not feel comfortable delegating a task to the agent.
- *Privacy and security*: Privacy and security of the user's data and profile are a concern for users. Questions such as how much a PDA can/should know about its user and what the control mechanisms are remain open.

Technical challenges

Technical challenges include

- *Experience scaling*: It is critical to integrate third-party applications and services into PDAs to scale both reactive and proactive experiences that can be handled by PDAs. Building the right tools and infrastructure to easily enable such integration is an open problem. User feedback shows that unsupported experiences are one of the single biggest sources of user dissatisfaction. For example, a user wants to be able book a taxi or calculate the mortgage payments for a house using speech with his or her PDA, but these, and many more scenarios, may not be supported by the PDA.
- *Speech recognition challenges*: Despite all the recent progress in speech recognition with the application of deep learning techniques, issues such as background noise, speaker accent, Bluetooth, side speech, pocket dial, and unintentional wake up remain to be addressed [9]. To truly fulfill its promise, a PDA should recognize all the personal words that the user cares about. This includes any name (not just English), any place, and any thing (e.g., user's contact list), essentially leading to an open-domain, unlimited vocabulary speech recognition problem.
- *LU challenges*: Domain scaling to cover many more domains, high-quality LU model development, and continual refinement with feedback loop data are the main challenges. Building reusable models across different tasks is an important problem to solve as well.
- *DM*: Heterogeneous knowledge back ends and application interfaces are a bottleneck for expanding the domains and tasks a PDA can cover. The APIs for different applications/services for the actions they perform as well as data/knowledge back ends are not standardized and require custom interface work and query building.
- *Locale/market expansion*: Building a proactive or reactive experience, not for English but for other languages and markets, is another open problem. This requires reusing or building all the resources (e.g., data, content, and models) and capabilities for new locales and markets.
- *Different device/end points*: Even though smartphones were the initial target device for deploying PDAs, soon it became evident that the underlying intelligence and agent capability should be highlighted in other devices and end points. For example, Cortana started with the phone, but it is now made available in PCs and tablets and even on Xbox. This in turn creates another problem; not all experiences make sense for a given device, or the same query may be interpreted differently on different devices. For example, users cannot send SMS from a PC or Xbox using PDAs, but they can do it on a phone. The query "go home" could mean "get driving directions to home" on the phone experience, but it may mean "go to shell" on the Xbox device.
- *Service challenges*: PDAs use numerous services to enable a given scenario. There are also software engineering and service challenges that impact the overall user experience. For example, if the latency for handling a user request is too high, it reflects negatively on the user experience. In fact, instability may even stop users from using the scenario altogether. Likewise, users expect high reliability and availability (e.g., > 99.9%) from the services handling the requests.

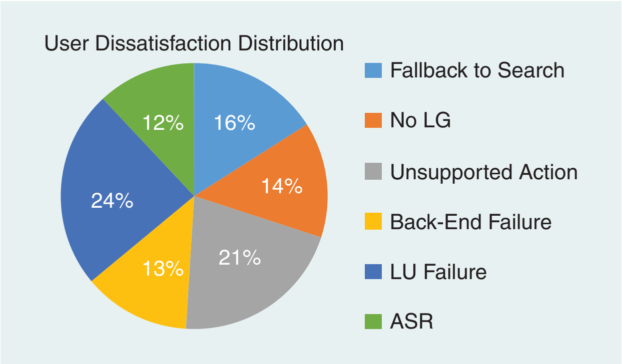


FIGURE 12. The E2E overall user dissatisfaction distribution over different components.

All of these are requirements and constraints that influence the system design.

Moving forward

Research on human work habits and task management [3], [63], [64] shows that people usually complete all their important tasks yet may fail to successfully complete tasks with soft deadlines or may forget less-critical details. In the short term, PDAs can provide great utility by becoming the digital memory that users can depend on for help with completion of everyday tasks. In fact, it is these scenarios that are used most by the users (e.g., reminders, meetings, and some proactive notifications and alerts).

Because time is a critical asset, improving personalized time management and utility through proactive and reactive task delegation and completion seems to be a plausible and desirable long-term goal for PDAs. In the future, it is the scalable and seamless third-party integration that can substantially increase the scenario and experience coverage and determine whether PDAs will fulfill the promise of a true personal assistant that users can depend upon to manage their personal and work life, effectively making them more productive. The walls between applications may start to break down if PDAs achieve app/service composition to complete new tasks in a scalable way.

PDAs will surface on many different devices and environments. This will create new signal processing challenges, such as accurate speaker separation and tracking in multispeaker environments (e.g., home, car), robustness with respect to different device types, and robust speech recognition across age, gender, and accent. Advances in algorithms, signal processing, and machine learning would be needed to solve these problems.

On the industry front, the investment and competition in PDA technology will keep increasing over the next decade. It is seen by some that PDAs may set the balance of power in the next phase of the Internet, if it becomes the gateway to applications and services with the proliferation of IoT devices. It is too early to call it an inflection point for PDA technology. It is likely that true natural language human–computer interaction with gadgets may take another decade to be second nature.

Acknowledgments

I would like to thank the past and present members of the Language Understanding and Dialog Systems Group at Microsoft, who built the conversational understanding and DM capabilities of Cortana: Alex Rochette, Asli Celikyilmaz, Beatriz Diaz Acosta, Chandra Akkiraju, Daniel Boies, Danko Panic, Derek Liu, Divya Jetley, Diamond Bishop, Elizabeth Krawczyk, Gabrielle Knight, Hisami Suzuki, Jean-Phillipe Robichaud, John Nave, Khushboo Aggarwal, Kjel Larsen, Logan Stromberg, Minwoo Jeong, Nikhil Ramesh, Omar Zia Khan, Paul Crook, Puyang Xu, Rachel Morton, Ravi Bikkula, Roman Holenstein, Roy Tan, Steve Kofsky, Tasos Anastasakos, Vasiliy Radostev, Vipul Agarwal, Young-Bum Kim, and Zhaleh Feizollahi.

Author

Ruhi Sarikaya (rsarikay@amazon.com) is the director of applied science at Amazon. He received his B.S. degree from Bilkent University, Ankara, Turkey, in 1995; his M.S. degree from Clemson University, South Carolina, in 1997; and his Ph.D. degree from Duke University Durham, North Carolina, in 2001, all in electrical and computer engineering. He was a principal science manager at Microsoft from 2011 to 2016, where he founded and managed the team that built language understanding and dialog management capabilities of Cortana and Xbox One. Before Microsoft, he was with IBM Research for ten years. Prior to joining IBM in 2001, he was a researcher at the University of Colorado at Boulder for two years. He has authored more than 100 technical papers and has more than 60 issued/pending patents. He is a Senior Member of the IEEE.

References

[1] T. Mitchell, R. Caruana, D. Freitag, J. McDermott, and D. Zabowski, "Experience with a learning personal assistant," *Commun. ACM*, vol. 37, no. 7, pp. 80–91, 1994.

[2] H. Chalupsky, Y. Gil, C. Knoblock, K. Lerman, J. Oh, D. Pynadath, T. Russ, and M. Tambe, "Electric elves: Agent technology for supporting human organizations," *AI Mag.*, vol. 23, no. 2, pp. 11–24, 2002.

[3] K. Myers, P. Berry, J. Blythe, K. Conley, M. Gervasio, D. McGuinness, D. Morley, A. Pfeffer, et al., "An intelligent personal assistant for task and time management," *AI Mag.*, vol. 28, no. 2, pp. 47–61, 2007.

[4] SRI International. (2003–2009). CALO: Cognitive Assistant that Learns and Organizes. [Online]. Available: pal.sri.com

[5] N. Yorke-Smith, S. Saadati, K. L. Myers, and D. N. Morley, "The design of a proactive personal agent for task management," *Int. J. Artif. Intell. Tools*, vol. 21, no. 1, 2012.

[6] S. Sohrab, M. Zhang, C. J. Karr, S. M. Schueller, M. E. Corden, K. P. Kording, and D. C. Mohr, "Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study," *J. Med. Internet Res.*, vol. 17, no. 7, p. e175, 2015.

[7] S. Schiaffino and A. Amandi, "User-interface agent interaction: Personalization issues," *Int. J. Human-Computer Studies*, vol. 60, no. 1, pp. 129–148, 2004.

[8] P. Maes, "Agents that reduce work and information overload," *J. ACM*, vol. 37, no. 7, pp. 30–40, 1994.

[9] J. Li, L. Deng, R. Haeb-Umbach, and Y. Gong, *Robust Automatic Speech Recognition, a Bridge to Practical Applications*. New York: Academic, 2015.

[10] G. Saon, H. K. J. Kuo, S. Rennie, and M. Picheny, "The IBM 2015 English conversational telephone speech recognition system," in *Proc. Interspeech*, Dresden, Germany, 2015.

[11] G. Tür and R. De Mori, Eds., *Spoken Language Understanding: Systems for Extracting Semantic Information from Speech*. Hoboken, NJ: Wiley, 2011.

[12] A. Deoras, R. Sarikaya, G. Tür, and D. Hakkani-Tür, "Joint decoding for speech recognition and semantic tagging," in *Proc. Interspeech*, 2012, pp. 1067–1070.

[13] R. Sarikaya, G. Hinton, and A. Deoras, "Application of deep belief networks for natural language understanding," *IEEE Trans. Audio, Speech, Language Process.*, vol. 22, no. 4, pp. 778–784, 2014.

[14] N. Gupta, G. Tür, D. Hakkani-Tür, S. Bangalore, G. Riccardi, and M. Gilbert, "The AT&T spoken language understanding system," *IEEE Trans. Audio, Speech, Language Process.*, vol. 14, no. 1, pp. 213–222, Jan. 2006.

[15] J. P. Robichaud, P. Crook, P. Xu, O. Z. Khan, and R. Sarikaya, "Hypotheses ranking for robust domain classification and tracking in dialogue systems," in *Proc. Interpseech*, Singapore, 2014, pp. 145–149.

[16] P. A. Crook, A. Marin, V. Agarwal, K. Aggarwal, T. Anastasakos, R. Bikkula, D. Boies, A. Celikyilmaz, S. Chandramohan, Z. Feizollahi, R. Holenstein, M. Jeong, O. Z. Khan, Y. B. Kim, E. Krawczyk, X. Liu, D. Panic, V. Radostev, N. Ramesh, J. P. Robichaud, A. Rochette, L. Stromberg, and R. Sarikaya, "Task completion platform: A self-serve multi-domain goal oriented dialogue platform," in *Proc. North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* San Diego, CA, 2016, pp. 47–51.

[17] J. R. Bellegarda, "Spoken language understanding for natural interaction: The Siri experience," in *Natural Interaction with Robots, Knowbots and Smartphones.* New York: Springer-Verlag, 2014, pp. 3–14.

[18] O. Z. Khan, J. P. Robichaud, P. Crook, and R. Sarikaya, "Hypotheses ranking and state tracking for a multi-domain dialog system using ASR results," in *Proc. Interspeech*, Dresden Germany, 2015.

[19] A. J. Hunt and A. W. Black. "Unit selection in a concatenative speech synthesis system using a large speech database," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, 1996, vol. 1, pp. 373–376.

[20] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, et al., "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups" *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, 2012.

[21] H. Sak, A. W. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in *Proc. Interspeech*, Singapore, 2014, pp. 338–342.

[22] A. Bhargava, A. Celikyilmaz, D. H. Tur, and R. Sarikaya, "Easy contextual intent prediction and slot detection," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, May 2013, pp. 8337–8341.

[23] P. Xu and R. Sarikaya, "Contextual domain classification in spoken language understanding systems using recurrent neural network," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, 2014, pp. 136–140.

[24] C. Liu, P. Xu, and R. Sarikaya, "Deep contextual language understanding in spoken dialogue systems," in *Proc. Interspeech*, 2015, pp. 120–124.

[25] R. Sarikaya, P. A. Crook, A. Marin, M. Jeong, J. P. Robichaud, A. Celikyilmaz, Y. B. Kim, A. Rochette, O. Z. Khan, X. Liu, D. Boies, T.

Anastasakos, Z. Feizollahi, N. Ramesh, H. Suzuki, R. Holenstein, E. Krawczyk, and V. Radostev, "An overview of end-to-end language understanding and dialog management for personal digital assistants," in *Proc. IEEE Spoken Language Technology,* San Diego, CA, 2016.

[26] T. Winograd, "Understanding natural language," *Cognitive Psych.*, vol. 3, no. 1, pp. 1–191, 1972.

[27] W. A. Woods, "Transition network grammars for natural language analysis," *Commun. ACM*, vol. 13, no. 10, pp. 591–606, 1970.

[28] W. Ward and S. Issar, "Recent improvements in the CMU spoken language understanding system," in *Proc. Workshop Human Language Technology, Association for Computational Linguistics*, 1994, pp. 213–216.

[29] R. F. Simmons, "Natural language question-answering systems: 1969," *Commun. ACM*, vol. 13, no. 1, pp. 15–30, 1970.

[30] T. Strzalkowski and S. Harabagiu, Eds., *Advances in Open-Domain Question-Answering.* Berlin, Germany: Springer-Verlag, 2006.

[31] C. L. Isbell and J. S. Pierce, "An IP continuum for adaptive interface design," in *Proc. Human–Computer Interaction Int. 2005*, Las Vegas, NV.

[32] E. Levin, R. Pieraccini, and W. Eckert, "A stochastic model of human-machine interaction for learning dialog strategies," *IEEE Speech Audio Process.*, vol. 8, no. 1, pp. 11–23, 2000.

[33] Y. Shi, Y. C. Pan, M. Y. Hwang, K. Yao, H. Chen, Y. Zou, and B. Peng, "A factorization network based method for multi-lingual domain classification," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, 2015, pp. 5276–5280.

[34] G. E. Hinton, S. Osindero, and T. Yee-Whye, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.

[35] I. H. Witten and F. Eibe, *Data Mining: Practical Machine Learning Tools and Techniques*. San Mateo, CA: Morgan Kaufmann, 2005.

[36] X. Liu and R. Sarikaya, "A model based approach to weight dictionary entities for spoken language understanding," in *Proc. IEEE Spoken Language Technology,* Lake Tahoe, NV, 2014.

[37] D. Ferrucci, E. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A. Kalyanpur, A. Lally, J. W. Murdock, et al., "Building Watson: An overview of the DeepQA project," *AI Mag.,* vol. 31, no. 3, pp. 59–79, 2010.

[38] D. Boies, R. Sarikaya, A. Rochette, Z. Feizollahi, and N. Ramesh, "Techniques for updating partial dialog state," U.S. Patent Application 20150095033*,* 2013.

[39] A. Celikyilmaz, Z. Feizollahi, D. Hakkani-Tür, and R. Sarikaya, "Resolving referring expressions in conversational dialogs for natural user interfaces," in *Proc. Empirical Methods on Natural Language Processing,* 2014, pp. 2094–2104.

[40] N. Shadbolt, W. Hall, and T. Berners-Lee, "The semantic web revisited," *IEEE Intell. Syst.*, vol. 21, no. 3, pp. 96–101, 2006.

[41] R. Baeza-Yates and A. Tiberi, "Extracting semantic relations from query logs," in *Proc. ACM Special Interest Group on Knowledge Discovery and Data Mining*, 2007, pp. 76–85.

[42] L. Heck, D. Hakkani-Tür, and G. Tür, "Leveraging knowledge graphs for web-scale unsupervised semantic parsing," in *Proc. Interspeech*, August 2013, pp. 1594–1598.

[43] A. El-Kahky, X. Liu, R. Sarikaya, G. Tür, D. Hakkani-Tür, and L. Heck, "Extending domain coverage of language understanding systems via intent transfer between domains using knowledge graphs and search query click logs," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Florence, Italy, 2014, pp. 4067–4071.

[44] Y. Ma, P. Crook, and R. Sarikaya, "Statistical inference over knowledge graphs for task-oriented spoken dialog systems," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Brisbane, Australia, 2015, pp. 5346–5350.

[45] Y.-B. Kim, M. Jeong, and R. Sarikaya, "Semi-supervised slot tagging with partially labeled sequences from web search click logs," in *Proc. North American Chapter of the Association for Computational Linguistics,* 2015, pp. 84–92.

[46] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor, "Freebase: A collaboratively created graph database for structuring human knowledge," in *Proc. 2008 ACM SIGMOD Int. Conf. Management of Data*, New York, 2008, pp. 1247–1250.

[47] S. Seneff, E. Hurley, R. Lau, C. Pao, P. Schmid, and V. Zue, "GALAXY-II: A reference architecture for conversational system development," in *Proc. Int. Conf. Spoken Language Processing,* 1998, vol. 98, pp. 931–934.

[48] P. S. Huang, X. He, J. Gao, L. Deng, A. Acero, and L. Heck, "Learning deep structured semantic models for web search using clickthrough data," in *Proc. 22nd ACM Int. Conf. Information and Knowledge Management*, 2013, pp. 2333–2338.

[49] V. Zue, S. Seneff, J. R. Glass, J. Polifroni, C. Pao, T. J. Hazen, and L. Hetherington, "Jupiter: A telephone-based conversational interface for weather information," *IEEE Speech Audio Process.*, vol. 8, no. 1, pp. 85–96, 2000.

[50] R. Pieraccini and J. Huerta, "Where do we go from here? Research and commercial spoken dialog systems," in *Proc. 6th SIGdial Workshop on Discourse and Dialogue*, 2005.

[51] Y. B. Kim, K. Stratos, R. Sarikaya, and M. Jeong, "New transfer learning techniques for disparate label sets," in *Proc. Association for Computational Linguistics,* 2015, pp. 473–482.

[52] J. Huang and E. N. Efthimiadis, "Analyzing and evaluating query reformulation strategies in web search logs," in *Proc. 18th ACM Conf. Information and Knowledge Management*, 2009, pp. 77–86.

[53] S. Young, M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, B. Thomson, and K. Yu, "The hidden information state model: A practical framework for POMDP-based spoken dialogue management," *Comput. Speech Lang.*, vol. 24, no. 2, pp. 150–174, 2009.

[54] A. Marin, P. Crook, O. Z. Khan, V. Radostev, K. Aggarwal, and R. Sarikaya, "Flexible, rapid authoring of goal-orientated, multi-turn dialogues using the task completion platform," in *Proc. Interspeech 2016*, San Francisco, CA, 2016, pp. 1571–1572.

[55] C. J. C. Burges, K. M. Svore, P. N. Bennett, A. Pastusiak, and Q. Wu, "Learning to rank using an ensemble of lambda-gradient models," *J. Mach. Learn. Res.*, vol. 14, pp. 25–35, 2011.

[56] W. T. Yih, X. He, and C. Meek. "Semantic parsing for single-relation question answering," in *Proc. Association for Computational Linguistics,* 2014, pp. 643–648.

[57] M. Nickel, M, K. Murphy, V. Tresp, and E. Gabrilovich, "A review of relational machine learning for knowledge graphs," *Proc. IEEE*, vol. 104, no. 1, pp. 11–33, 2016.

[58] M. Dinarelli, A. Moschitti, and G. Riccardi, "Discriminative reranking for spoken language understanding," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 2, pp. 526–539, Feb. 2012.

[59] V. Rieser and O. Lemon, "Natural language generation as planning under uncertainty for spoken dialogue systems," in *Empirical Methods in Natural Language Generation*. Berlin, Germany: Springer-Verlag, 2010, pp. 105–120.

[60] R. J. Kate, Y. W. Wong, and R. J. Mooney, "Learning to transform natural to formal languages," in *Proc. 12th Nat. Conf. Artificial Intelligence (AAAI-05)*, Pittsburgh, PA, 2005, pp. 1062–1068.

[61] A. H. Oh and A. I. Rudnicky, "Stochastic language generation for spoken dialogue systems," in *Proc. 2000 ANLP/NAACL Workshop on Conversational Systems*, 2000, vol. 3, pp. 27–32.

[62] J. Jiang, A. H. Awadallah, R. Jones, U. Ozertem, I. Zitouni, R. G. Kulkarni, and O. Z. Khan, "Automatic online evaluation of intelligent assistants," in *Proc. 24th Int. World Wide Web Conf.*, 2015, pp. 506–516.

[63] M. Czerwinski, E. Horvitz, and S. Wilhite, "A diary study of task switching and interruptions," in *Proc. ACM Conf. Human Factors in Computing Systems,* Vienna, Austria, 2004, pp. 175–182.

[64] V. Bellotti, B. Dalal, N. Good, P. Flynn, D. G. Bobrow, and N. Ducheneaut, "What a to-do: Studies of task management towards the design of a personal task list manager," in *Proc. ACM Conf. Human Factors in Computing Systems,* Vienna, Austria, 2004, pp. 735–742.

[65] Statista. (2016). Number of mobile phone users worldwide from 2013 to 2019 (in billions). [Online]. Available: http://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/

[66] Siri. *Apple*. [Online]. Available: http://www.apple.com/ios/siri/

[67] Google Now. *Google*. [Online]. Available: https://www.google.com/landing/now/

[68] Cortana. *Microsoft*. [Online]. Available: https://www.microsoft.com/en-us/mobile/experiences/cortana/

[69] Alexa. *Amazon Developer*. [Online]. Available: https://developer.amazon.com/public/solutions/alexa

[70] M. Meeker. (2016, June 1). Internet trends 2016—Code conference. Kleiner Perkins Caufield Byers. Menlo Park, CA. [Online]. Available: http://www.kpcb.com/blog/2016-internet-trends-report

[71] Samsung. (2014, Apr. 23). 10 sensors of Galaxy S5: Heart rate, finger scanner and more. *Samsung Newsroom*. [Online]. Available: http://news.samsung.com/global/10-sensors-of-galaxy-s5-heart-rate-finger-scanner-and-more

SP

Hana Godrich

Students' Design Project Series: Sharing Experiences

The fast pace of technology refresh offers new opportunities in engineering training and education. Undergraduate engineering programs are seeking the right balance between theory and practice, as students are increasingly expected to have more knowledge of cutting-edge technologies along with the fundamental understanding of engineering concepts and modeling tools. Instructors constantly look for opportunities to enrich the learning experience through demonstrations and hands-on experience to address a growing demand from the industry for engineers with "know-how" skills. This can be accomplished through traditional channels such as instructional laboratories, independent study, and research activities or, alternatively, through student design project programs. In the United States, many universities and colleges have programs for senior design projects in place, also referred to as capstone projects. These engineering projects are an assimilation of knowledge and capabilities built over the first three years in undergraduate studies and, in many cases, are deeply reliant on signal and information processing expertise.

A background in signal and information processing is fundamental to many engineering applications, yet it is gradually built through a sequence of engineering courses that bring students in their senior year to a position where it can be applied successfully. Low-cost, multifeature computer platforms, being developed at an outstanding rate, alongside open-source software, offer implementation capabilities like never before. The availability of high-performance, low-cost, and small footprint single-board computers and microcontrollers systems presents versatile building blocks that can be used by the students based on their individual interests and applied in multidisciplinary projects. As an example, Raspberry Pi [1] was introduced to the market in 2012 and has since gone through three generations with the latest being Raspberry Pi 3B. In a mere 8.56 × 5.65-cm package, a weight of 45 g, and a price tag of US$35, Raspberry Pi 3B offers a 1.2-GHz 64-bit Quad Core ARM Cortex-A53 central processing unit (CPU) with 1 GB of memory, on-board network Ethernet, wireless and Bluetooth, 17 GPIO general-purpose input/output (GPIO) and more. A smaller version of it was introduced in May 2016 in the form of Raspberry Pi Zero, offering a 1-GHz CPU with 512 MB of memory and 40 GPIO for US$5 in a 6.5 cm × 3 cm package and a weight of 9 g. Intel's Edison computer module [2], a wide range of Arduino products [3], and Texas Instruments [4] introduce similar opportunities, to name a few. These popular products are accompanied by large selection of add-ons, built-in libraries, interplatform computability, and open source software that support the development of diverse, signal and information processing-based applications in the Internet of things, wearables, three-dimensional printing, autonomous vehicles, drones, biomedical devices, and more.

Embedded systems such as smartphones, smart watches, and health-tracking wristbands, are part of our daily lives and, at times, offer a more narrow application use and limited resources. They are commonly used in telecom/datacom, automation, military, medical, and automotive applications. Some companies open their systems to the research and development of new applications. Google released the source code for Google Glass [5], declaring it as an open-to-hackers platform. Apple provides access to its sensory system and enables turning an iPhone into a medical diagnostic device through Apple Research Kit [6]. iRobot released the Create 2 Programmable robot [7], an open-sourced electronics prototyping platform that enables integration of sensory systems and microcomputers. These are a few examples in a growing trend with exciting possibilities for students and developers by gaining access to hardware and software tools that can be manipulated for varies applications in computer vision, robotics, machine learning, cybersecurity, biomedical, and biometrics, to name a few.

This progress in affordable and accessible hardware and software tools can

Digital Object Identifier 10.1109/MSP.2016.2620157
Date of publication: 11 January 2017

 1053-5888/17©2017IEEE

be instrumental in enhancing students' hands-on experience during undergraduate studies while building solid theoretical foundations. A new article series and contributions platform has been developed by a team of guest editors for *IEEE Signal Processing Magazine* who wish to promote the sharing of experiences and best practices with undergraduate studies, engineering projects, research, and innovation in diverse signal and information processing applications. Students and advisers have been invited to contribute information and practices in engineering design projects in SigPort (https://sigport.org/events/spm-student-design-project-series#citation-ieee). The goal of this initiative is to start a discussion on the role of experimental and project-based practices in modern signal and data processing education and support fast-track progress by sharing "know-how" experience. The SigPort submissions give a first glimpse into the potential of this idea. With more than 400 downloads within a time span of just a few weeks, it seems that there is a need within the signal and information processing community that should be further explored and advanced.

The projects submitted as the first wave of response to the call for contributions articulate the field's diversity and the multitude of applications, methods, and hardware and software tools explored by students and faculty.

One of these projects, "Graph Frequency Analysis of Brain Signals," by Leah Goldsberry, Weiyu Huang, and Dr. Alejandro Ribeiro with the Department of Electrical and Systems Engineering in the University of Pennsylvania, Philadelphia (http://arxiv.org/pdf/1512.00037v2.pdf), is evaluating the practice of graph signal processing methods in neuroscience. Motivated by the need to identify effective means to introduce the relationship between data and information, commonly achieved through the use of alternative representation of data, the project is designed to explore the relationship between neuroscience and graph Fourier transform (GFT). The notion of the GFT and graph filters is presented to decompose a given subject's brain signal into sections that represent

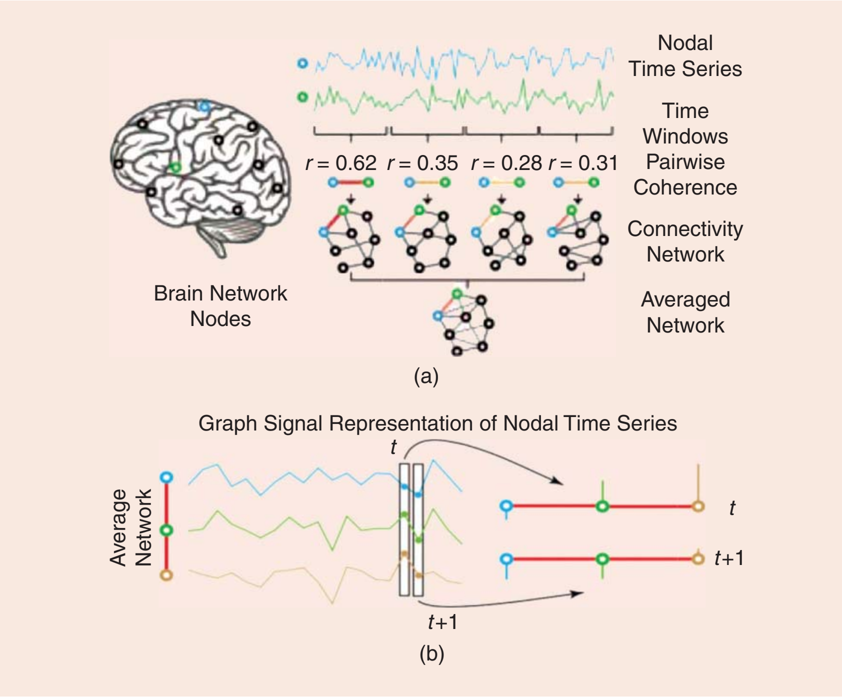


FIGURE 1. (a) A brain network representation using average functional coherence values. (b) A brain graph signal using regional fMRI data for each time point t. See the original figure and report at https://sigport.org/documents/graph-frequency-analysis-brain-signals.

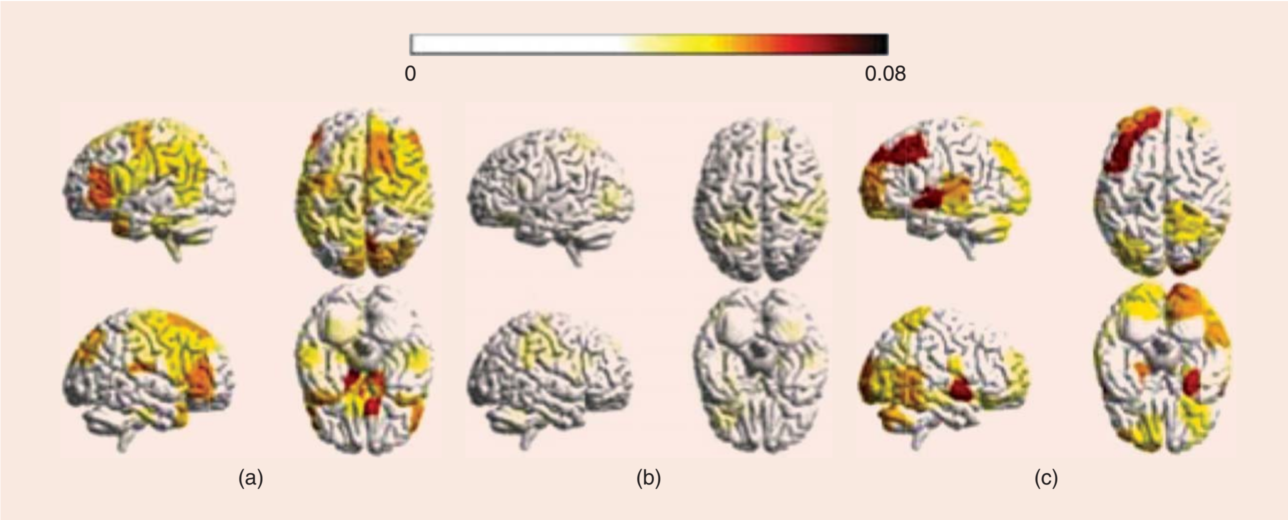


FIGURE 2. An example distribution of decomposed signals across all brain regions for the first experiment. Average energy with respect to (a) x_L, (b) x_M, and (c) x_H. A thresholding is applied.

various modes of variability. The student may then autonomously explore different frequencies or temporal variability. Functional magnetic resonance imaging (fMRI)-based experiments studied the response to visual cues over a training period, aiding in the modeling of average connectivity in brain networks. Filtering is then used to decompose the graph signals to look for a correlation between the decomposed signals and a subject's performance when learning a task (see Figures 1 and 2).

Hand prosthesis controls, based on electromyography (EMG) signals, are addressed in a separate project, "Real-Time Control of Hand Prosthesis Using EMG," by Or Dicker, Aviv Peleg, Dr. Tal Shnitzer, Dr. Oscar Lichtenstein, and Dr. Yair Moshe with the Signal and Image Processing Lab, Andrew and Erna Viterbi Faculty of Electrical Engineering, Technion–Israel Institute of Technology, Haifa, Israel (https://sigport.org/documents/real-time-control-hand-prosthesis-using-emg). This project is motivated by the need to design a low-cost multifunctional alternative for high-cost prostheses for below-the-elbow amputees. It sets a clear scope of work in terms of reliability, portability, gestures support, interface quality, and setup time along with a low price target. The Intel Edison board is used along with EMG sensors (Myo armband) and a printed prostatic hand. The project includes data collection from the EMG sensors sampled at 200 Hz and communicated to the microcomputer via Bluetooth. Data is processed to determine the required hand gesture via the implementation of simple feature extraction at each time segment. The mean absolute value method was found to perform best for this case. Each time segment was classified to one of six gesture classes using K-nearest neighbors. The prototype was tested to demonstrate high classification success rates and multiple gestures support with a low cost (evaluated at US$345). Figures 3–5 illustrate the design and functionalities.

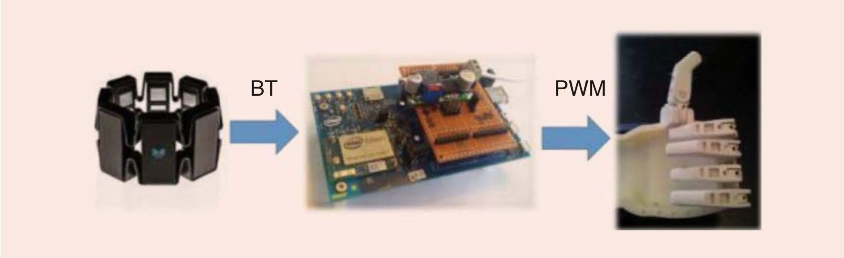


FIGURE 3. An overview of the project "Real-Time Control of Hand Prosthesis Using EMG System."

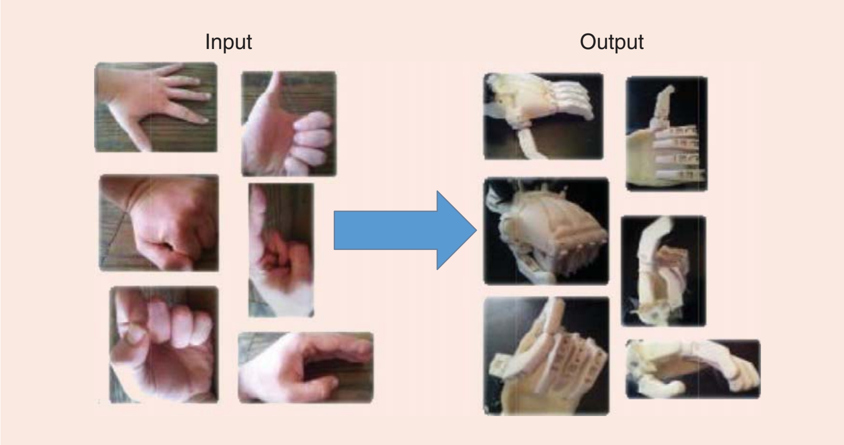


FIGURE 4. The six implemented hand gestures.

FIGURE 5. The constructed prosthetic hand.

Another closely related project, "Micro Hand Gesture Recognition System Using Ultrasonic Active Sensing Method," by Yu Sang, Quan Wang, and Dr. Yimin Liu with the Intelligent Sensing Lab, Department of Electronic Engineering, Tsinghua University, Beijing, China (https://www.youtube.com/watch?v=8FgdiIb9WqY; https://sigport.org/documents/micro-hand-gesture-recognition-system-using-ultrasonic-active-sensing-method), looks into the use of ultrasonic active sensing methods for microhand gesture recognition. The pulsed radar signal processing technique is used to obtain time-sequential range-Doppler features. Object distance and velocity are measured through a single channel to reduce hardware complexity. A hidden Markov model approach is used to classify time-sequential range-Doppler features. A state transition mechanism significantly compresses the data and extract intrinsic signatures. A real-time prototype was developed and an average recognition accuracy of 90.5% for seven gestures was achieved. Related works include the WiSee gesture recognition system developed in 2013 by Patel et al. [8] to leverage wireless signals for home sensing and recognition under complex conditions and Google's Soli project that

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

uses wearable and micro hand gestures to control smart devices [9]. Figures 6–8 illustrate the design and functionalities.

A tutorial for a do-it-yourself Sky Imager is offered in "DIY Sky Imager for Weather Observation," by Soumyabrata Dev (Nanyang Technological University), Florian M. Savoy (University of Illinois at Urbana Champaign's Singapore and Advanced Digital Sciences Center), Dr. Yee Hui Lee (Nanyang Technological University), and Dr. Stefan Winkler (University of Illinois at Urbana Champaign's Singapore and Advanced Digital Sciences Center); for more information see https://github.com/FSavoy/DIY-sky-imager and https://sigport.org/documents/diy-sky-imager-weather-observation.

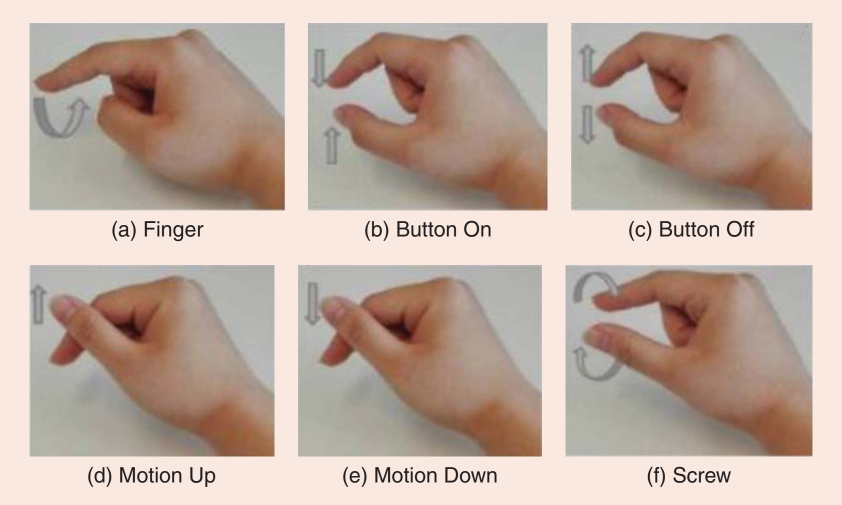


FIGURE 6. Examples of micro hand gestures, named by (a) finger, (b) button on (BtnOn), (c) button off (BtnOff), (d) motion up (MtnUp), (e) motion down (MtnDn), and (f) screw.

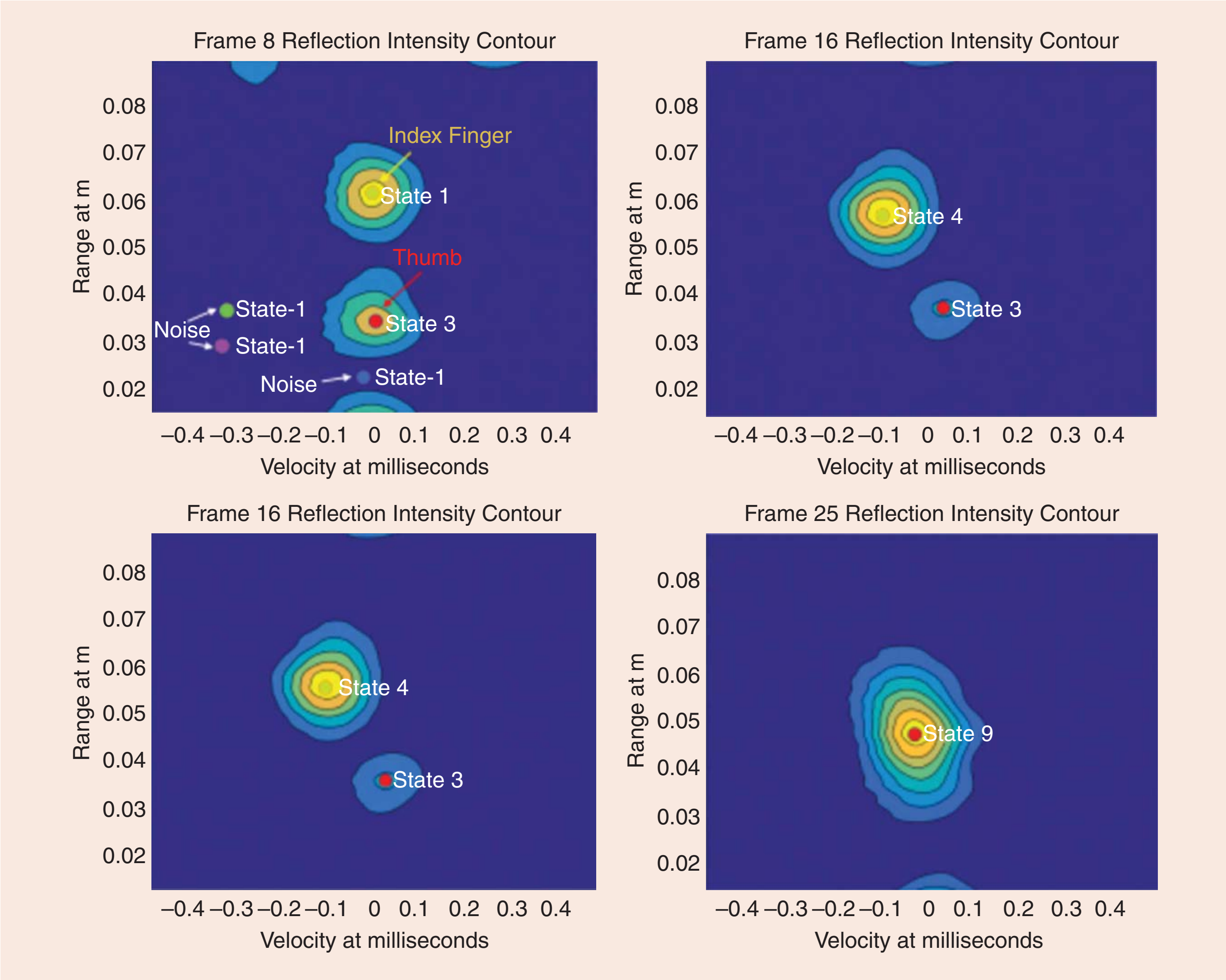


FIGURE 7. Range-Doppler feature frames sampled in the "button off" (see Figure 1) gesture. (a) Index finger and thumb separate at a 3-cm distance. Noises will be removed using time smoothing to increase robustness as the three marked "noise" objects in the first subfigure. (b) and (c) The index finger is moving down with acceleration while the thumb almost keeps static with a tiny velocity moving up. (d) Two fingers get touched. Note that the object's trajectory will always be a curve in the range-Doppler plane. The symbolized states are labeled at the center of the detected objects.

As shown in Figure 9, this image processing application is Raspberry Pi-based and incorporates off-the-shelf components to replace what would otherwise be a high-cost system. The project involved a digital single-lens reflex camera and image processing performed by the microcomputer to generate a segmented binary image. A simple algorithm is employed to calculate the instantaneous cloud coverage. This is useful for regular monitoring of cloud formation over a region.

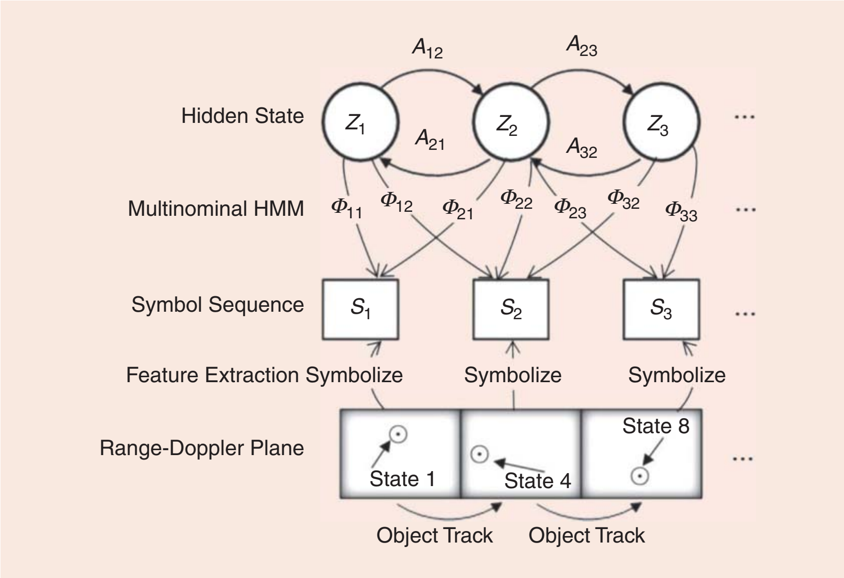


FIGURE 8. Feature extraction and classification workflow.

FIGURE 9. The design of the DIY Sky Imager.

Chien-Sheng Yang and Dr. Lav R. Varshney with the University of Illinois at Urbana-Champaign present a more research-oriented project in "Self-Sustainable OFDM Transmissions with Smooth Energy Delivery" (https://sigport.org/documents/self-sustainable-ofdm-transmissions-smooth-energy-delivery). This project studies the question: "Is it possible to have small peak-to-average power ratio (PAPR) in the cyclic prefix of an OFDM signal, while maintaining self- sustainability?" A new system architecture, shown in Figure 10, is proposed that employs a frame-theoretic method that demonstrates significant improvement in PAPR of the cyclic prefix in self-sustainable OFDM.

Audio processing is at the core of the "Automatic Lyrics Display System for Live Music Performances" project by Karan Vombatkere, Bochen Li, and Dr. Zhiyao Duan with the University of Rochester, New York (https://sigport.org/documents/automatic-lyrics-display-system-live-music-performances). The primary objective of the project is to design and implement a computational system that can follow live music performances (e.g., choruses) in real time and display pre-encoded lyrics for

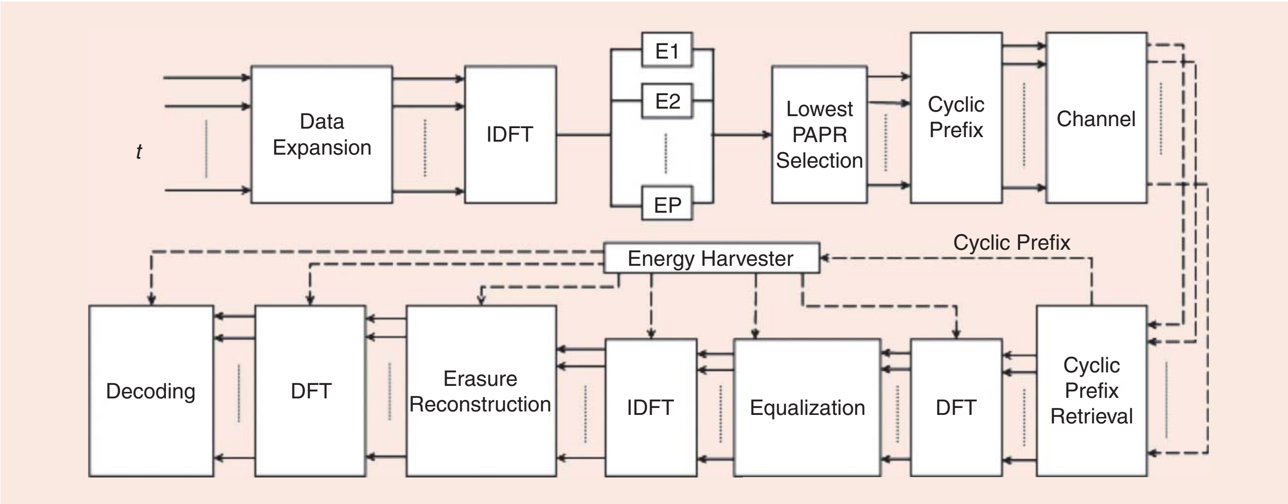


FIGURE 10. A self-sustainable OFDM system with EPS-CP for PAPRCP reduction.

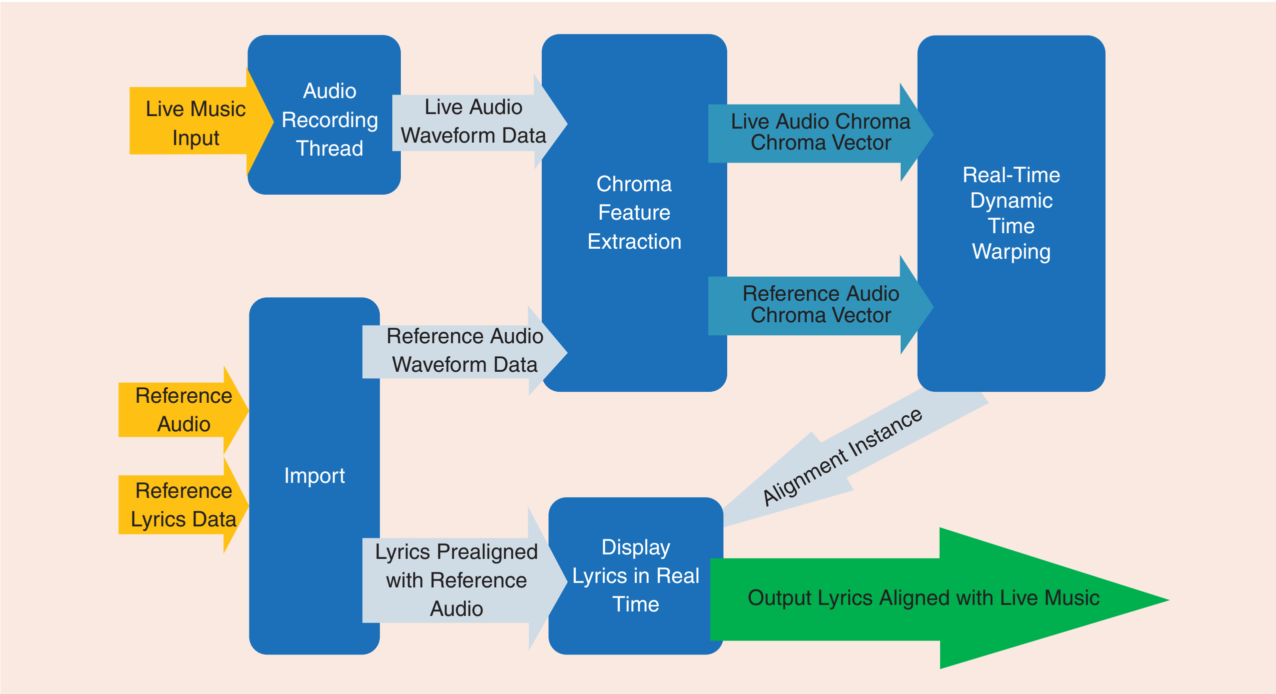


FIGURE 11. Lyrics display system flowchart.

the audience. As basic harmonic progression of two recordings is similar, the chroma feature is used to represent the audio data. It uses 12 bins to represent the relative energy of the audio in the 12 semitones of a musical octave, represents the harmonic content of the audio. A prerecorded version of a concert serves as reference for lyrics alignment. This is performed through online dynamic time wrapping. See the design and interface shown in Figures 11 and 12, respectively.

A National Instruments (NI)-based project is detailed in "Acoustic Detection and Localization of Impulsive Events in Urban Environments" by Sabeeh Irfan Ahmad, Hassan Shahbaz, Hassam Noor, Dr. Momin Uppal, and Dr. Muhammad Tahir with the Syed Babar Ali School of Science and Engineering, Lahore University of Management Sciences, Pakistan. The project focused on the detection of impulsive acoustic events and localizing the source in an urban environment. As illustrated in Figure 13, the system used an array of microphones that recorded the sounds and transmitted raw data to a central fusion center, based on the NI compact reconfigurable input-output setup. With a sampling frequency of 100 KHz per channel and on-board field-programmable gate array unit programmed to minimize latency, the collected data was process to estimate the source angle-of-arrival. The sine sweep method was found to be the most effective, where a sinusoid of temporally increasing frequency, both linearly and exponentially, is applied via accurate speakers.

This first set of project highlights manifest a diverse collection of undergraduate engineering projects and tools. Some concentrate on developing devices that offer a solution for social or environmental needs. Others are focused on mathematical modeling and research to advance existing technology. A tutorial project offered step-by-step guidelines for building an image acquisition and processing platform. Many projects concentrated on embedded systems and all incorporated signal and information processing methods commonly thought in fundamental courses.

The main objective of this effort is to expand the knowledge base in the

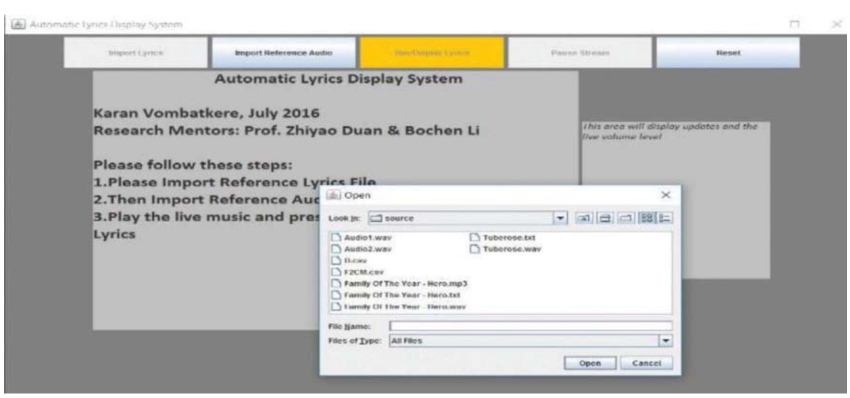


FIGURE 12. A graphic user interface.

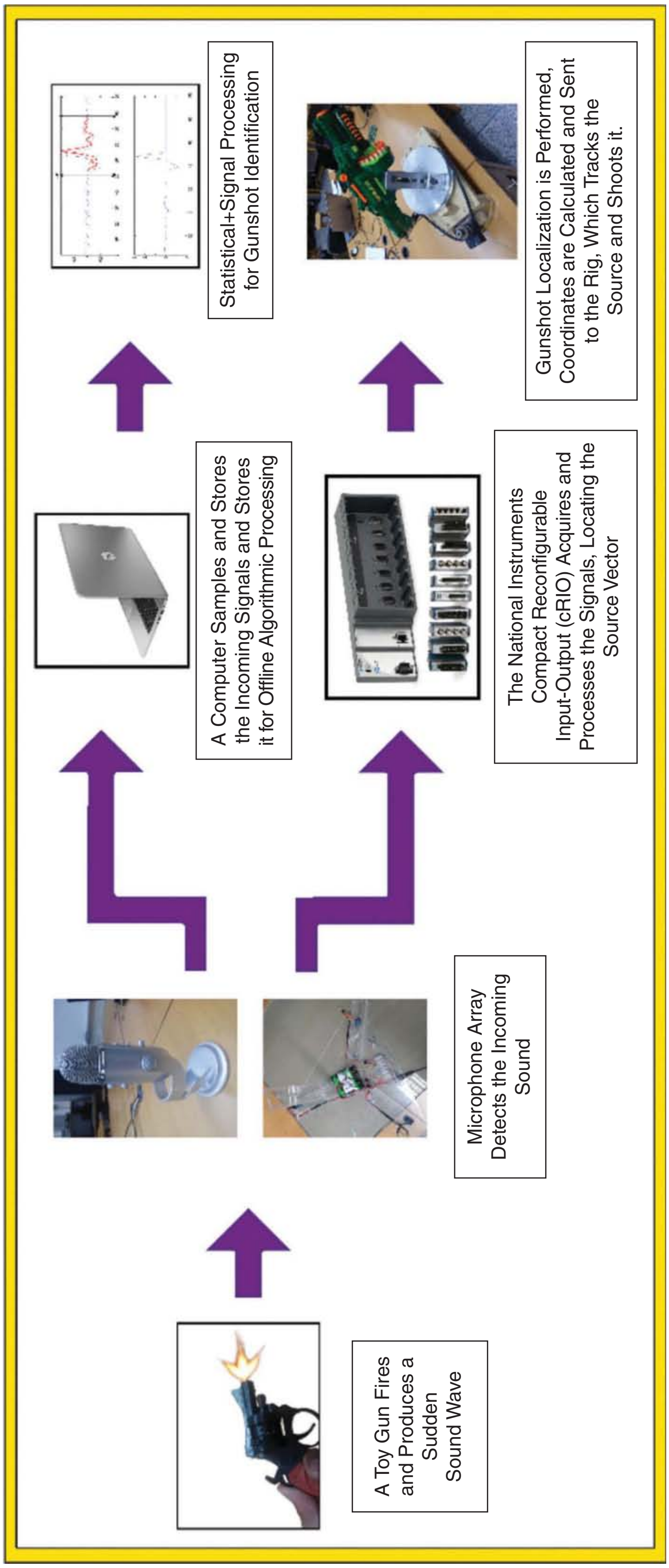


FIGURE 13. The test bed for data acquisition and offline processing.

community and shorten the learning curve in the academia while encouraging a discussion in the community on education and engineering studies allowing for rethinking teaching method to a promote more disruptive and multidisciplinary engineering.

Author

Hana Godrich (godrich@soe.rutgers.edu) received the B.Sc. degree in electrical engineering from the Technion Israel Institute of Technology, Haifa, in 1987; the M.Sc. degree in electrical engineering from Ben-Gurion University, Beer-Sheva, Israel, in 1993; and the Ph.D. degrees in electrical engineering from the New Jersey Institute of Technology, Newark, in 2010. She is the undergraduate program director in the Electrical and Computer Engineering department at Rutgers University, Piscataway, New Jersey. Her research interests include statistical signal processing with application to wireless sensor networks, communication, smart grid, and radar systems. She is a Senior Member of the IEEE.

References

[1] RaspberryPi. [Online]. Available: https://www.raspberrypi.org/

[2] RaspberryPi. [Online]. Available: https://github.com/raspberrypi/documentation/blob/master/setup/README.md

[3] Arduino. [Online]. Available: https://www.arduino.cc/en/Guide/Windows;

[4] Texas Instruments, Inc. [Online]. Available: http://www.ti.com/

[5] Google.com. [Online]. Available: https://developers.google.com/glass/

[6] Apple.com. [Online]. Available: http://www.apple.com/researchkit/

[7] iRobot.com. [Online]. Available: http://www.irobot.com/About-iRobot/STEM/Create-2/Projects.aspx

[8] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proc. 19th Annu. Int. Conf. Mobile Computing & Networking*, ACM, 2013, pp. 27–38.

[9] Google ATAP. (May 29, 2016). Project soli [Online]. Available: https://atap.google.com/soli/2015

SP

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

Denis A. Gudovskiy and Lichung Chu

An Accurate and Stable Sliding DFT Computed by a Modified CIC Filter

The sliding discrete Fourier transform (SDFT) is a popular algorithm used in nonparametric spectrum estimation when only a few frequency bins of an M-point discrete Fourier transform (DFT) are of interest. Although the classical SDFT algorithm described in [1] is computationally efficient, its recursive structure suffers from accumulation and rounding errors, which lead to potential instabilities or inaccurate output. Duda [2] proposed a modulated SDFT (mSDFT) algorithm, which has the property of being guaranteed stable without sacrificing accuracy, unlike previous approaches described in [1], [3], and [4]. However, all of these conventional SDFT methods presume DFT computation on a sample-by-sample basis. This is not computationally efficient when the DFT needs only to be computed every $R\ (R > 1)$ samples. To address such cases when R-times downsampling is needed, Park et al. [5] proposed a hopping SDFT (HDFT) algorithm. Recently, Wang et al. [6] presented a modulated HDFT (mHDFT) algorithm, which combines the HDFT algorithm with the mSDFT idea to maintain stability and accuracy at the same time. In parallel, Park [7] updated the HDFT algorithm with its guaranteed stable modification called gSDFT, which exists only for certain M and L relationships.

Digital Object Identifier 10.1109/MSP.2016.2620198
Date of publication: 11 January 2017

Introduction

In this article, we present a method to improve the SDFT algorithm. The improved algorithm is called *cascade integrator-comb* (*CIC*)-*SDFT*, which is based on extending the idea established in the mSDFT algorithm and mostly in the context of spectrum estimation application. Similarly to mSDFT, we move the DFT bin of interest k to the zero position to exclude complex coefficient multiplication in the recursive stage and avoid instabilities. In addition, CIC-SDFT comprises a modified CIC filter structure proposed by Hogenauer [8]. Two goals are achieved using this approach. First, the accuracy of spectrum estimation is improved by using high-order CIC filters without computationally expensive windowing in the frequency domain as described in [1]. Second, the complexity of the SDFT can be further decreased by reducing the DFT output rate, also achieved by [5]–[7].

SDFT and mSDFT

The kth frequency bin of an M-point DFT at time index n for input signal x is defined by

$$X_n(k) = \sum_{m=0}^{M-1} W_M^{-km} x_{q+m}, \qquad (1)$$

where $q = n - M + 1,\ 0 \le k \le M - 1$, and the complex exponential factor $W_M = e^{j2\pi/M}$. A recursive equivalent of (1) is given by

$$X_n(k) = W_M^k \left(X_{n-1}(k) + x_n - x_{n-M} \right). \qquad (2)$$

Equation (2) can be implemented as a filter with a comb stage followed by an integrator stage, as shown in Figure 1(a). This conventional SDFT filter has a z-domain pole on the unit circle located at $z = W_M^k$. Hence, it is only marginally stable in finite precision recursive calculations with accumulation, except at points when poles $z = \pm 1$ or $z = \pm j$.

Duda [2] proposed to shift the $X(k)$ DFT bin of interest to the $k = 0$ bin prior to calculating the comb stage in SDFT. Thus, the $X_n(k)$ calculation is simplified to multiplication of input signal x_n by the modulation sequence W_M^{-km}, followed by calculation of a new zero-frequency $Y_n(0)$ DFT bin expressed as

$$X_n(k) = Y_n(0) = Y_{n-1}(0) + y_n - y_{n-M}, \qquad (3)$$

where $y_n = W_M^{-km} x_n$ and $y_{n-M} = W_M^{-k(m-M)} x_{n-M}$.

The mSDFT structure is shown in Figure 1(b). Complex multiplication in the integrator stage is unnecessary because $W_M^0 = 1$. Therefore, the mSDFT filter becomes guaranteed-stable and accurate at the same time. In addition, complex multiplication in the recursive stage might limit the clock rate of the digital circuit, which is avoided in this method by effectively moving it into the feedforward part.

The drawback of the mSDFT algorithm compared to the conventional

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

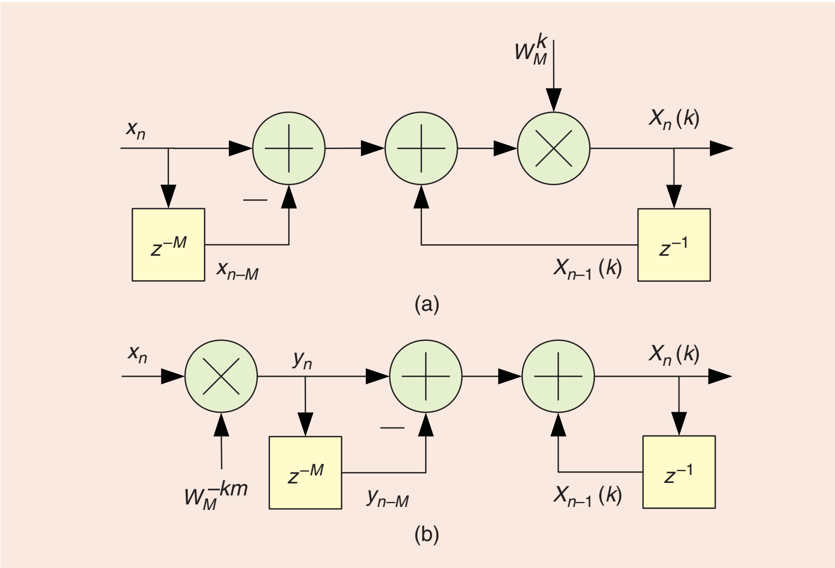


FIGURE 1. SDFT filter structures: (a) conventional and (b) mSDFT.

SDFT is the necessity to generate a modulating sequence W_M^{-km} instead of keeping a fixed twiddle factor W_M^k. Note that the output of the mSDFT in Figure 1(b) has $W_M^{-k(m+1)}$ phase shift compared to DFT, which does not have an effect on magnitude spectrum estimation applications. The last two topics are not covered in this article since they are well described in [2].

CIC-SDFT

Recursive structure

One can note that the mSDFT filter structure depicted in Figure 1(b) looks exactly like a modified first-order CIC filter with an additional complex multiplication by W_M^{-km}. This is not a coincidence, since the mSDFT calculates the $Y_n(0)$ (zero frequency) DFT bin given by

$$Y_n(0) = \sum_{m=0}^{M-1} y_{q+m}. \tag{4}$$

Equation (3) rewrites the moving average filter (4) in recursive form, which, indeed, is a first-order CIC filter without rate change. Using (3) and (4), we can generalize the mSDFT idea and apply the CIC filter theory [8] for DFT spectrum estimation.

The general structure of CIC-SDFT is depicted in Figure 2, which contains complex multiplication of input signal x_n by the modulating sequence W_M^{-km}, followed by a CIC decimation filter with R rate change. The CIC decimator contains an integrator section with L integrator stages, a downsampler by R, and a comb section with L comb stages. The CIC filter part is equivalent to a cascade of L moving average filters with transfer function expressed as

$$H(z) = \frac{(1-z^{-M})^L}{(1-z^{-1})^L} = \left(\sum_{m=0}^{M-1} z^{-m}\right)^L. \tag{5}$$

It is known that the magnitude response of the CIC filter evaluated at $z = e^{j\omega}$ can be written as

$$|H(\omega)| = \left|\frac{\sin(M\omega/2)}{\sin(\omega/2)}\right|^L, \tag{6}$$

where ω is a normalized angular frequency and $-\pi \le \omega \le \pi$ radians/sample. In this case, the magnitude response $|H(\omega)|$ of the CIC filter is the magnitude response $W(\omega)$ of a window function $w(m)$ applied to a discrete-time Fourier transform (DTFT), when a finite length DFT is being computed. Note that the computed CIC-SDFT magnitude response should be normalized according to the CIC decimator gain, which is equal to M^L.

Window function

The described CIC-SDFT algorithm provides two salient features. First, it improves spectrum estimation performance using a naturally embedded B-spline window function $w(m)$, which is defined as self-convolution of L length-M rectangular functions. Equation (6) shows that, for $L = 1$, CIC-SDFT provides an exact DFT spectrum. However, the spectral leakage can be reduced by increasing the CIC filter order L, which is equivalent to a higher-order filter magnitude response $W(\omega)$. For example, it can be shown that $W_{L=1}(\omega)$ corresponds to a rectangular window of length M in the time domain, $W_{L=2}(\omega)$ corresponds to a triangular window of length $(2M-1)$, and so on.

In general, windowing is an expensive operation. Conventional time-domain windowing would compromise computational simplicity of the SDFT algorithm. Hence, frequency-domain convolution of adjacent DFT outputs with another window function was proposed in [1]. Practically, it is limited to only short window functions with preferably power-of-two coefficients because SDFT complexity grows faster than a linear function of window length. For comparison, Figure 3 illustrates several DFT magnitude responses for $M = 32$. First, it shows

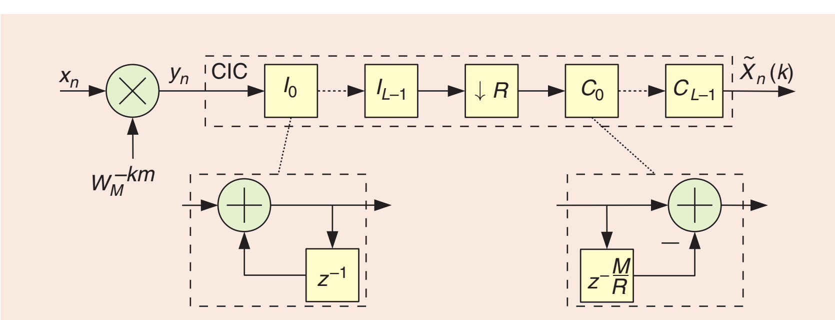


FIGURE 2. The recursive CIC-SDFT filter structure.

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

three normalized magnitude responses of $W_{L=1,2,3}(\omega)$ window functions generated by CIC-SDFT with lengths M, $(2M-1)$, and $(3M-2)$, respectively. The last plot depicts the magnitude response of the optimized five-point Hanning window described in [1], which has three nonzero coefficients –1/4, 1/2, and –1/4. As can be seen, CIC-SDFT realizes a powerful window function compared to the short Hanning window. For example, an interference that falls into the first sidelobe will be attenuated by 13, 26, and 39 dB for $L = 1, 2, 3$, respectively.

Estimator variance

Assuming that the input signal x is corrupted by additive white Gaussian noise (AWGN) with variance $\sigma^2 = 1/\mathrm{SNR}$, where SNR is the signal-to-noise ratio, the CIC-SDFT variance can be written as

$$\operatorname{var}\{\hat{X}_n(k)\} = C/\mathrm{SNR}, \qquad (7)$$

where coefficient C depends on the convolution of L rectangular windows. Each rectangular window can be represented as a length-M column vector of all ones. Then, coefficient C can be expressed as

$$C = \frac{1}{M^{2L}} \sum_{m=0}^{L(M-1)+1} \left(\underbrace{\mathbf{1} * \mathbf{1} * \ldots * \mathbf{1}}_{L}\right)_m^2. \qquad (8)$$

The closed-form expression of (8) can be written in vector form as

$$C = \| \mathbf{A}^{L-1}\mathbf{b} \|^2 / M^{2L}, \qquad (9)$$

where vector $\mathbf{b} = \begin{bmatrix}\mathbf{1}\\ \mathbf{0}\end{bmatrix}$, $\mathbf{1}$ is an $M \times 1$ column vector of all ones, $\mathbf{0}$ is an $(L-1)(M-1) \times 1$ column vector of all zeros, and $\mathbf{A}$ is a Toeplitz matrix with $\mathbf{b}$ as the first column and $1 \times (L(M-1)+1)$ vector $\mathbf{u} = [1\ 0\ 0 \ldots 0]$ as the first row. The case $L = 1$ gives exactly the variance of a moving average filter $\sigma^2_{L=1} = 1/(M\,\mathrm{SNR})$. The variance of a second-order filter is decreased by a factor of 3/2 for any large M. The variance of the output periodogram for high SNR can be approximated by

$$\operatorname{var}\{|\hat{X}_n(k)|^2\} \approx 2C/\mathrm{SNR}. \qquad (10)$$

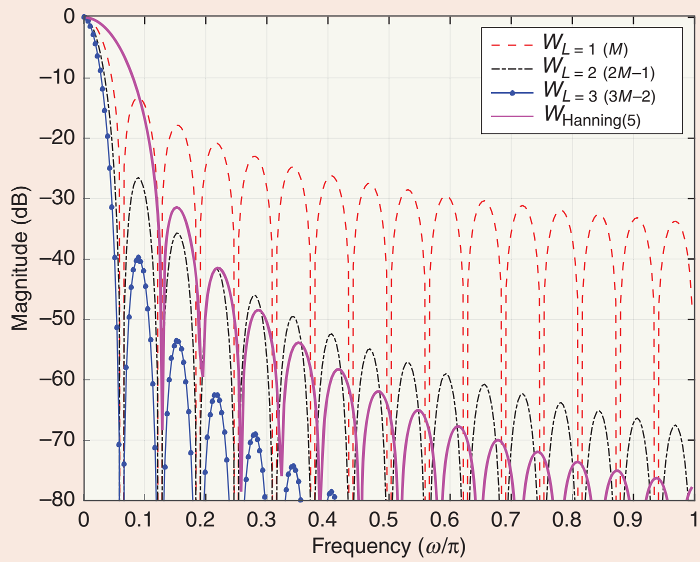


FIGURE 3. The window function comparison.

Computational complexity

The second feature of CIC-SDFT is reduced computational complexity for $R > 1$ cases. When $R = 1$ and $L = 1$, CIC-SDFT performs the same number of computations as mSDFT. When $R > 1$, the digital circuit after the integrator section operates at a f_s/R clock rate and the memory size in the comb section is decreased by a factor of R. That is a significant reduction in computational complexity and considerably simplifies digital circuit implementation. Table 1 summarizes computational complexity by comparing the number of complex multiplications, complex additions, and memory size needed to calculate a single-bin DFT.

In addition, it states whether a particular method requires generation of the modulating sequence and the output phase shift correction or not, which can be accomplished by a number of approaches. For example, a simple table-based method can be used for some applications to generate the complex exponent W_M^{-km} or, in the general case, a generator and a phase shift corrector described in [2] can be used. The phase shift corrector is not needed for magnitude spectrum estimation. Note that a conventional SDFT in Table 1 is the method that cannot be guaranteed stable and accurate at the same time due to recursive complex multiplication. The gSDFT and mHDFT computational workload was recalculated for the single-bin case. As can be seen, these algorithms are not beneficial in this configuration unless a significant subset of M bins has to be calculated. On the other hand, CIC-SDFT is a beneficial method, when only one or a few bins of an M-point DFT need to be computed with R downsampling.

Table 1. Single-bin DFT computational complexity.

Method	Mul.	Add.	Mem.	Mod. Seq.	Ph. Cor.
DFT	M	$M-1$	0	No	No
SDFT	1	2	$M+1$	No	No
mSDFT	1	2	$M+1$	Yes	Yes
gSDFT, mHDFT	1	2	$M+2R-1$	No	No
CIC-SDFT	1	$L\left(1+\frac{1}{R}\right)$	$L\left(1+\frac{M}{R}\right)$	Yes	Yes

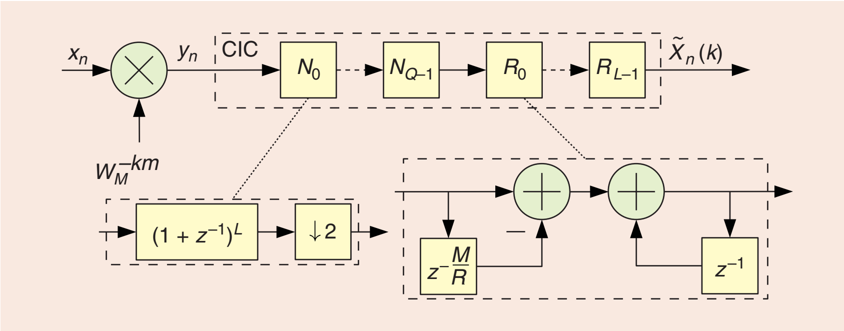


FIGURE 4. The partially nonrecursive CIC-SDFT filter structure.

Partially nonrecursive structure

Although recursive CIC-SDFT reduces computational complexity for $R > 1$, it experiences a bit-growth according to CIC filter theory. Assuming an input bit-width B_{in}, the bit-width B used for all computations in the CIC-SDFT can be expressed as

$$B = \lceil L\log_2(M)\rceil + B_{\text{in}}. \tag{11}$$

Wider bit-widths have to be used as the number of stages L and DFT length M grows. This drawback can be solved by using nonrecursive structures derived from polynomial factoring and applying polyphase decomposition [9]. Then for power-of-two DFT length $M = 2^P$, the transfer function of CIC-SDFT can be written as

$$\begin{aligned} H(z) &= \left(\sum_{m=0}^{M-1} z^{-m}\right)^L \\ &= (1+z^{-1})^L(1+z^{-2})^L \\ &\quad \times (1+z^{-4})^L \ldots (1+z^{-2^{P-1}})^L. \end{aligned} \tag{12}$$

Assuming a power-of-two downsampling factor $R = 2^Q$, (12) can be rewritten as

$$\begin{aligned} &H(z) \\ &= \underbrace{(1+z^{-1})^L(1+z^{-2})^L\cdots(1+z^{-2^{Q-1}})^L}_{H_1(z)} \\ &\quad \times \underbrace{(1+z^{-2^Q})^L\cdots(1+z^{-2^{P-1}})^L}_{H_2(z)} \\ &= H_1(z)H_2(z). \end{aligned} \tag{13}$$

Next, assuming that the $H_1(z)$ output is followed by a downsampler by R, the transfer function $H_2(z)$ in (13) can be simplified to

$$\begin{aligned} H_2(z) &= (1+z^{-1})^L \ldots (1+z^{-2^{P-Q-1}})^L \\ &= \left(\sum_{m=0}^{M/R-1} z^{-m}\right)^L. \end{aligned} \tag{14}$$

From (13) and (14) it is clear that the CIC-SDFT transfer function $H(z)$ can be split into two parts: $H_1(z)$ computed in nonrecursive fashion and $H_2(z)$ computed in recursive fashion.

The first nonrecursive part comprises Q stages where each stage increases bit-width by L bits and downsamples the output by two. For each stage, which calculates $(1+z^{-1})^L$, several implementations are possible. For example, it can be realized as a length-L cascade of $(1+z^{-1})$ operations or direct exponentiation of the whole stage. The latter can be expressed as a transfer function $H_N(z)$, and for $L = 4$

$$\begin{aligned} H_N(z) &= (1+z^{-1})^4 \\ &= 1 + 6z^{-2} + z^{-4} \\ &\quad + 4z^{-1}(1+z^{-2}) \\ &= H_{N_1}(z) + H_{N_2}(z), \end{aligned} \tag{15}$$

where $H_{N_1}(z)$ and $H_{N_2}(z)$ are new polyphase components. Note that each polyphase component may downsample computations by two prior to performing add operations. The nonrecursive part experiences bit growth of only $L\log_2(R)$, which means that the total bit-width at its output can be written as

$$B_N = L\log_2(R) + B_{\text{in}}. \tag{16}$$

The second recursive part now implements an Lth-order moving-average filter of length M/R with the transfer function $H_2(z)$. Such a filter can be realized using a cascade of L sections, where each section contains a comb stage followed by an integrator stage. Then, the bit-width increases by only $\log_2(M/R) = (P - Q)$ bits per section. The total bit-width at the output of partially nonrecursive structure is identical to (11).

Figure 4 illustrates an alternative partially nonrecursive implementation of CIC-SDFT. First, it contains Q nonrecursive stages $N_0, \ldots N_{Q-1}$ according to the aforementioned description. Second, nonrecursive stages are followed by a cascade of L recursive sections $R_0, \ldots R_{L-1}$. Due to the fact that bit-widths are increased on a per-stage basis rather than all at once at the input of the CIC-SDFT, the computational and circuit complexity decreases. Moreover, since downsampling is performed as early as possible, the number of add operations is minimized as well. Another important advantage of the structure shown in Figure 4 is the ability to model CIC-SDFT using floating-point arithmetic, since overflows in the integrator stages are avoided. The latter property allows to implement a rounding operation between algorithm stages.

Summary

In this article, a novel SDFT algorithm called CIC-SDFT was introduced. It generalizes a previous approach by incorporating a modified CIC filter structure. Such a generalization adds two new programmable parameters: filter order and output rate. Filter order is responsible for an embedded window function, and therefore determines the spectrum estimator variance and interference rejection capabilities. Sidelobe level is proportional to the filter order as $-13L$ dB for the embedded window function. The closed-form expression for the variance of CIC-SDFT estimator and its periodogram was provided.

Programmability of the output rate allows one to decrease algorithm computational complexity when needed. Specifically, the number of memory cells and half of the add operations in recursive CIC-SDFT are inversely proportional to the downsampling factor R.

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

Bit growth can be minimized using the presented partially nonrecursive structure of CIC-SDFT, which is suitable for digital circuit implementation and algorithm modeling using floating-point arithmetic.

Acknowledgments

We would like to thank Dennis R. Morgan for improving this manuscript.

Authors

Denis A. Gudovskiy (denis.gudovskiy@us.panasonic.com) is a design engineer with Panasonic, Cupertino, California.

Lichung Chu (lichung.chu@olympus.com) is a director of engineering with Olympus Communication Technology of America, San Diego, California.

References

[1] E. Jacobsen and R. Lyons, "The sliding DFT," *IEEE Signal Process. Mag.*, vol. 20, no. 2, pp. 74–80, Mar. 2003.

[2] K. Duda, "Accurate, guaranteed stable, sliding discrete Fourier transform," *IEEE Signal Process. Mag.*, vol. 27, no. 6, pp. 124–127, Nov. 2010.

[3] E. Jacobsen and R. Lyons, "An update to the sliding DFT," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 110–111, Jan. 2004.

[4] S. Douglas and J. Soh, "A numerically-stable sliding-window estimator and its application to adaptive filters," in *Proc. 31st Annual Asilomar Conf. on Signals, Systems, and Computers*, vol. 1, Nov. 1997, pp. 111–115.

[5] C. Park and S. Ko, "The hopping discrete Fourier transform," *IEEE Signal Process. Mag.*, vol. 31, no. 2, pp. 135–139, Mar. 2014.

[6] Q. Wang, X. Yan, and K. Qin, "High-precision, permanently stable, modulated hopping discrete Fourier transform," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 748–751, June 2015.

[7] C.-S. Park, "Fast, accurate, and guaranteed stable sliding discrete Fourier transform," *IEEE Signal Process. Mag.*, vol. 32, no. 4, pp. 145–156, July 2015.

[8] E. Hogenauer, "An economical class of digital filters for decimation and interpolation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 2, pp. 155–162, Apr. 1981.

[9] R. Losada and R. Lyons, "Reducing CIC filter complexity," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 124–126, July 2006.

SP

ERRATA

The author order in the article "Smart Driver Monitoring: When Signal Processing Meets Human Factors" in the November 2016 issue of *IEEE Signal Processing Magazine* [1] printed incorrectly due to a production error. We apologize for any confusion this may have caused. The correct order of the byline is as follows:

Amirhossein S. Aghaei, Huei-Yen Winnie Chen, George Liu, Cheng Liu, Zohreh Sojoudi, Dengbo He, Birsen Donmez, and Konstantinos N. Plataniotis.

Digital Object Identifier 10.1109/MSP.2016.2636256
Date of publication: 11 January 2017

Reference

[1] A. S. Aghaei, B. Donmez, C. C. Liu, D. He, G. Liu, K. N. Plataniotis, H. Y. Winnie Chen, and Z. Sojoudi, "Smart driver monitoring: When signal processing meets human factors," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 35–48, Nov. 2016.

SP

LECTURE NOTES

S.Y. Kung

Compressive Privacy: From Information/Estimation Theory to Machine Learning

Most of our daily activities are now moving online in the big data era, with more than 25 billion devices already connected to the Internet, to possibly over a trillion in a decade. However, big data also bears a connotation of "big brother" when personal information (such as sales transactions) is being ubiquitously collected, stored, and circulated around the Internet, often without the data owner's knowledge. Consequently, a new paradigm known as *online privacy* or *Internet privacy* is becoming a major concern regarding the privacy of personal and sensitive data.

As depicted in Figure 1, Internet data live in two different worlds: 1) the private sphere, where data owners generate and process decrypted data, and 2) the public sphere, where the data in cloud servers are presumably encrypted and therefore unaccessible by intruders. However, the data may be decrypted by the intended and trusted "authorities," who will be provided with the right key for the data decryption. Following the cryptographic channel and adversarial models, formally defined by Claude Shannon, we shall also name the data owner, intended user, and intruder as Alice, Bob, and Eve, respectively. Recall that the typical security protocol hinges upon Alice's passing a decrypting key to Bob but not to Eve. Since the notion of an unbreakable key is questionable, there is no wonder that Internet data remain highly vulnerable to unauthorized leakages and hacker attacks.

Digital Object Identifier 10.1109/MSP.2016.2616720
Date of publication: 11 January 2017

Data owner should have control over data privacy

Privacy-preserving (PP) tools have a broad spectrum of applications, covering numerous types of Internet data (such as speech, image, location, and media/social/health data). They all require a delicate balance between utilization and privacy. For example, in case of a bomb explosion, images from various mobile sources near the crime scene may be collected by authorities for wide-scale forensic analysis. Ideally, the uploaded images should provide critical and relevant information to help capture the targeted suspects while protecting the full facial images of the innocent from being leaked to the public.

Data are not just a collection of words/numbers working in isolation, rather they encompass the global and

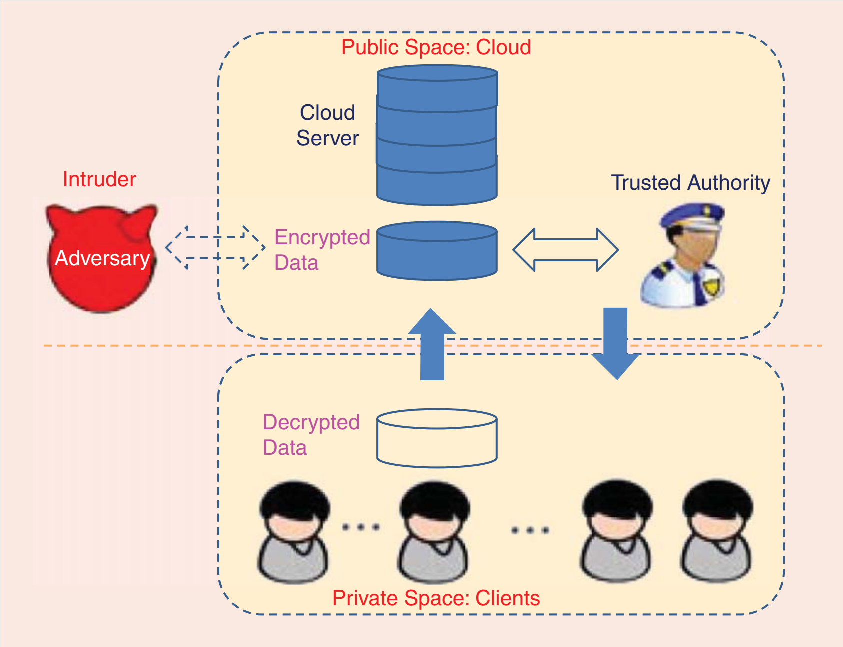


FIGURE 1. In collaborative learning environments, individual data are uploaded to the cloud. From the privacy perspective, data in the private sphere versus the public sphere should be treated differentially, which calls for a novel PP encoding paradigm, known as *compressive privacy* (*CP*). For privacy protection, the query data uploaded to the public sphere should be designed to retain the information useful for the intended application and should not be easily repurposed into malicious exploitation.

 1053-5888/17©2017IEEE

highly coordinated control of information. PP protocols allow the data owner to control the fate of the data, instead of chancing with the protection promised by the cloud severs. To this end, the Defense Advanced Research Projects Agency (DARPA) has championed a major Brandeis Program to develop novel communication protocols so that the uploaded data are useful only for the intended utility but not easily repurposed into privacy intrusion.

Privacy communication paradigm

Note that, as depicted in Figure 2, both Bob and Eve will receive the same data (denoted by $\mathbf{y}$), i.e., there is no key required. Ideally, in this paradigm, the query $\mathbf{y}$ should be useful to (friendly) Bob but useless to (malicious) Eve. More realistically, we would like to see to it that the query may retain information as lossless as possible to Bob and, at the same time, be as lossy as possible against Eve.

In short, it should also be recognized that the security protocols for data sharing are neither necessary nor sufficient for data privacy protection. Therefore, it is worth installing both the security and privacy protocols for maximal protection of Internet data.

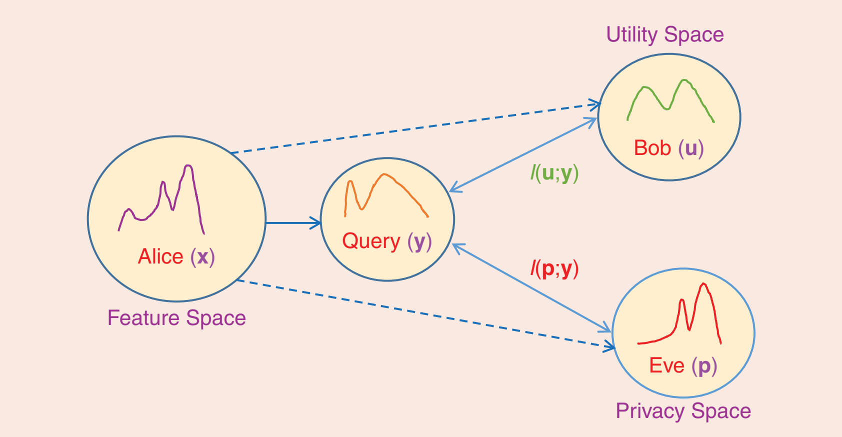


FIGURE 2. Our study on privacy preservation involves joint optimization over three design spaces: 1) the feature space (Alice), 2) the utility space (Bob), and 3) the privacy space (Eve). Alice (the data owner) wants to convey certain information relevant to Bob (the intended user/utilizer) while preventing it from being eavesdropped by Eve (the intruder). In the CP paradigm, both Bob and Eve will receive the same data (denoted by $\mathbf{y}$), i.e., there is no key required. We propose a query encoding scheme which is 1) information preserving from the utility's perspective but 2) information lossy from the perspective of privacy. Collectively, such a scheme is called CP. This article explores the utility-privacy tradeoff analysis via comparing the $I(\mathbf{u};\mathbf{y})$ and $I(\mathbf{p};\mathbf{y})$.

Differential privacy and compressive privacy

Differential privacy

In the differential privacy (DP) theory [1], a desensitizing function K is said to provide ϵ-DP of the data if $Pr[K(D) \in S] \leq e^{\epsilon} Pr[K(D') \in S]$ for all $S \in \text{Range}(K)$ and all data sets D and D' differing by one entry. Note also that the DP sensitization does not necessarily require the utility function to be known in advance, and DP guarantees that the distribution of the search result should be indistinguishable (modulo a factor of e^{ϵ}) with or without the missing entry. Two types of "differential-log-likelihood" criteria, ϵ-DP versus ϵ-information privacy (a stronger privacy metric), are analyzed and compared in the study of the so-called privacy funnel in [2]. Due to the absence of systematical methods for the derivation of optimal queries, however, the DP approach remains somewhat unwieldy for many real-world applications. This prompts us to explore other desensitizing methods for privacy preservation.

Compressive privacy

The query, denoted by $\mathbf{y}$, is represented by $\mathbf{y} = f(\mathbf{x}, \varepsilon)$ where $\mathbf{x}$ denotes the feature vector representing the original data and ε is an independent random noise. Unlike DP, the CP approach allows the query to be tailor-designed according to the known utility and privacy models. As depicted in Figure 2, we propose a query encoding scheme that is 1) information preserving from the utility's perspective but 2) information lossy from the privacy's perspective. Collectively, such a scheme is called *CP*. This article explores the tradeoff analysis between the utility mutual information (between $\mathbf{y}$ and Bob) and its privacy counterpart (between $\mathbf{y}$ and Eve). This is in a sharp contrast to most machine-learning problems, where the design goal is exclusively focused only on the utility information.

Scope and prerequisite of the article

This article explores the rich synergy between information theory, estimation theory, and machine learning and, ultimately, develops a PP methodology—CP. While formal courses on information theory, estimation theory, and machine learning are highly recommended for advanced and serious researchers, we shall nevertheless review the basic materials for novice readers, hopefully making the article somewhat self-contained.

Information and estimation theory

Let the original data (owned by Alice) be represented by a vector space containing M-dimensional random vectors

$$\mathbf{x} = [x_1, x_2, \ldots, x_M]^T.$$

To convey information concerning $\mathbf{x}$, the design of PP query $\mathbf{y}$ must be based on joint consideration of both the utility maximization (for Bob) and the privacy protection (against Eve). Mathematically,

- **Utility function:** The utility function is denoted by $\mathbf{u}(\mathbf{x})$.
- **Privacy function:** The privacy function (i.e., cost function) is denoted by $\mathbf{p}(\mathbf{x})$.

Prior work on non-Gaussian models

A natural formulation for the utility-privacy tradeoff analysis involves optimizing $I(\mathbf{u};\mathbf{y})$ [respectively, $I(\mathbf{p};\mathbf{y})$] while setting a bound on $I(\mathbf{p};\mathbf{y})$ [respectively, $I(\mathbf{u};\mathbf{y})$]. More specifically:

- **Information bottleneck (IB)** [4]. In the IB scenarios, it is assumed that

$\mathbf{p}(\mathbf{x}) = \mathbf{x}$, i.e., Eve has no additional role here. Two specific examples are as follows:

- Alice wants to transmit as much information as possible, relevant to Bob, while consuming minimum resources on bandwidth/storage.
- Alice wants to transmit as much information as possible, relevant to Bob, while preventing the original data from being reconstructed or leaked.

The objective is to design the query $\mathbf{y}$, which maximizes the utility gain, prescribed by $\mathbf{u}(\mathbf{x})$, while keeping the data bandwidth below a certain bound.

- **Privacy funnel (PF)** [2]. In this scenario, it is assumed that $\mathbf{u}(\mathbf{x}) = \mathbf{x}$, i.e., Bob has no role here. Now Alice wants to transmit the original data, as much as possible, while preventing leaking sensitive data to Eve. (Recall that Eve may perform any adversarial inference attack, based on the query $\mathbf{y}$, to intrude the privacy prescribed by $\mathbf{p}(\mathbf{x})$.) Therefore, we must design a query $\mathbf{y}$ that can minimize the privacy leakage while assuring a guaranteed level of information on the original data is being conveyed.

Most analysis regarding IB [4] and PF [2], especially those pertaining to the convex optimization, may be naturally extended to the general case when $\mathbf{u}(\mathbf{x}) \neq \mathbf{x}$ and $\mathbf{p}(\mathbf{x}) \neq \mathbf{x}$, i.e., both Bob and Eve have their own specific goals.

Gauss–Markov estimation theorem

The assumption of Gaussian distribution of the data is vital to our development of the CP theory. More specifically, it allows us to make use of the results that 1) the amount of information can be quantified as a log function of its variance and 2) the difference of variances (before-and-after-query) can be derived via the classic Gauss–Markov estimation theorem [3]. As proven next, the Gaussian assumption leads us to a simple (eigenvalue-based) optimal solution in closed form.

Now, as depicted in Figure 2, the design will be optimized over two competing linear vector spaces:

- **Utility subspace**. The utility function is represented by $\mathbf{u} = \mathbf{U}^T\mathbf{x}$, where $\mathbf{U} \in \Re^{M\times\mu}$ is the projection matrix characterizing the utility subspace.
- **Privacy subspace**. The cost (i.e., privacy) function is represented by $\mathbf{p} = \mathbf{P}^T\mathbf{x}$, where $\mathbf{P} \in \Re^{M\times\nu}$ is the matrix characterizing the privacy subspace. (Here we shall simply assume the subspace projection matrices $\mathbf{U}$ and $\mathbf{P}$ as given, leaving their learning/estimation strategies a later discussion.)

Gaussian distribution with linear optimal query

The linear query is represented by

$$y = \mathbf{f}^T\mathbf{x} + \varepsilon, \qquad (1)$$

where we assume that $\|\mathbf{f}\| = 1$ and ε is an independent random noise, with variance $\sigma_\varepsilon^2 = \sigma^2$, all without loss of generality.

The amount of information contained in $\mathbf{x}$ (as well as $\mathbf{u}$ and $\mathbf{p}$) can be quantitatively measured by its entropy

$$H(\mathbf{x}) \equiv -\int p(\mathbf{x}) \ \log p(\mathbf{x}) \ d\mathbf{x},$$

where the integration is taken over the M-dimensional vector space. Assume that $\mathbf{x}$ has a Gaussian distribution $N(\hat{\mathbf{x}}_0, \Sigma_x)$, where $\hat{\mathbf{x}}_0$ and $\Sigma_{\mathbf{x}}$ are the mean and covariance matrix of $\mathbf{x}$. The entropy and covariance matrix are closely connected

$$H(\mathbf{x}) = \frac{1}{2}\log_2(|\Sigma_{\mathbf{x}}|) + \frac{M}{2}\log_2(2\pi e).$$

Effect of query on covariance matrix

Assume that $E[\mathbf{x}] = \hat{\mathbf{x}}_0$ and $E[(\mathbf{x}-\hat{\mathbf{x}}_0)(\mathbf{x}-\hat{\mathbf{x}}_0)^T] = \Sigma_{\mathbf{x}}$. Before query, the initial estimate of $\mathbf{x}$ is $\hat{\mathbf{x}}_0$, and the initial error-covariance-matrix (ECM) is $E[(\mathbf{x}-\hat{\mathbf{x}}_0)(\mathbf{x}-\hat{\mathbf{x}}_0)^T] = \Sigma_{\mathbf{x}}$. Now that we are given the knowlege of the query, $y = \mathbf{f}^T\mathbf{x} + \varepsilon$. According to the well-known Gauss–Markov theorem; cf. [3, ch. 15], the optimal estimation of $\mathbf{x}$ is

$$\begin{aligned}\hat{\mathbf{x}} &= E[\mathbf{x} \mid \mathbf{y}]\\ &= \hat{\mathbf{x}}_0 + \sigma^{-2}(\Sigma_{\mathbf{x}}^{-1} + \sigma^{-2}\mathbf{f}\,\mathbf{f}^T)^{-1}\mathbf{f}[y - \mathbf{f}^T\hat{\mathbf{x}}_0].\end{aligned}$$

Let $\tilde{\mathbf{x}} \equiv \hat{\mathbf{x}} - \mathbf{x}$ denote the updated estimation error (given the query $\mathbf{y}$) with its corresponding ECM as follows:

$$\begin{aligned}\Sigma_{\tilde{\mathbf{x}}} &= E[\tilde{\mathbf{x}}\,\tilde{\mathbf{x}}^T \mid \mathbf{y}]\\ &= (\Sigma_{\mathbf{x}}^{-1} + \sigma^{-2}\ \mathbf{f}\ \mathbf{f}^T)^{-1}\\ &= \Sigma_{\mathbf{x}} - \Sigma_{\mathbf{x}}\mathbf{f}[\sigma^2 + \mathbf{f}^T\Sigma_{\mathbf{x}}\mathbf{f}]^{-1}\mathbf{f}^T\Sigma_{\mathbf{x}}.\end{aligned} \qquad (2)$$

This leads to a postquery Gaussian distribution for $\mathbf{x}$, denoted by $N(\hat{\mathbf{x}}, \Sigma_{\tilde{\mathbf{x}}})$. For multiquery cases, just change $\mathbf{f} \to \mathbf{F} \in \Re^{M\times m}$ and $\sigma^2 \to \Sigma_\epsilon \in \Re^{m\times m}$.

Effect on the estimation error of the original feature vector

Conventionally, we would like to maximally preserve the fidelity, i.e., to best reconstruct the original data. In this case, natural formulation for the optimal query vector(s) is as follows:

$$\text{argmax}_{\mathbf{f}}\ \{\text{trace}(\Sigma_{\tilde{\mathbf{x}}})\} \qquad (3)$$

whose optimal solution lies exactly on the principal component analysis (PCA) eigen-subspace.

Effect on the utility and privacy entropies

It is obvious that the additional query knowledge can only reduce the entropy of $\mathbf{u}$ and $\mathbf{p}$. The utility and privacy functions are linear functions of the state vector $\mathbf{x}$, so they are Gaussian distributed with the utility covariance matrices (before and after the query) given as

$$\Sigma_{\mathbf{u}} \equiv \mathbf{U}^T\Sigma_{\mathbf{x}}\mathbf{U} \text{ and } \Sigma_{\tilde{\mathbf{u}}} = \mathbf{U}^T\Sigma_{\tilde{\mathbf{x}}}\mathbf{U} \qquad (4)$$

and the privacy covariance matrices as

$$\Sigma_{\mathbf{p}} \equiv \mathbf{P}^T\Sigma_{\mathbf{x}}\mathbf{P} \text{ and } \Sigma_{\tilde{\mathbf{p}}} \equiv \mathbf{P}^T\Sigma_{\tilde{\mathbf{x}}}\mathbf{P}. \qquad (5)$$

Example 1. The double income problem (DIP)

Here, a two-dimensional vector $\mathbf{x} = [x_1\ x_2]^T$ represents the two individual incomes of a couple.

- From the utility perspective, to assess the family's total income, the utility function should be set as $\mathbf{u} = \mathbf{u}(\mathbf{x}) = x_1 + x_2$.
- From the privacy perspective, the query should not pry into the income disparity between the couple. To protect such privacy, the privacy function is set as $\mathbf{p} = \mathbf{p}(\mathbf{x}) = x_1 - x_2$.

Suppose that the initial covariance matrix of $\mathbf{x}$ is

$$\Sigma_{\mathbf{x}} = \begin{bmatrix} 8 & -6 \\ -6 & 10 \end{bmatrix}.$$

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

Note that $\mathbf{U} = [1\ 1]^T$ and $\mathbf{P} = [1 - 1]^T$ and the initial utility and privacy variances are respectively $\sigma_{\tilde{\mathbf{u}}}^2 = 6$ and $\sigma_{\tilde{\mathbf{p}}}^2 = 30$, cf. (4) and (5).

Let us study two scalar queries: $y = 0.8\,x_1 + 0.6\,x_2 + \varepsilon$ and $y' = 0.6\,x_1 + 0.8\,x_2 + \varepsilon$. According to (16), y and y' deliver almost the same differential information gain ≈ 2.7. However, because the spouse's income is initially more private than the husband's (note $10 > 8$), y and y' have different effects on the variances of $\mathbf{u}$ and $\mathbf{p}$.

In Figure 3, the distribution of $\mathbf{p}, \tilde{\mathbf{p}}$, and $\tilde{\mathbf{p}}'$ are respectively represented by the solid green, dotted red, and dashed blue curves. More specifically, y' is more intrusive than y, implying a higher weight placed on the spouse's income would leak more on the disparity. Indeed, we have $\sigma_{\tilde{\mathbf{p}}}^2 = 29.35$ and $\sigma_{\tilde{\mathbf{p}}'}^2 = 25.72$, cf. (5). On the other hand, from the utility perspective, y' outperforms y in producing a narrower utility variance: $\sigma_{\tilde{\mathbf{u}}'}^2 = 1.72 < \sigma_{\tilde{\mathbf{u}}}^2 = 1.96$; cf. (4).

□

Utility mutual information maximization

In information theory, $H(\mathbf{u})$ denotes the entropy of $\mathbf{u}$, $I(\mathbf{u};\mathbf{y})$ denotes the mutual information between $\mathbf{u}$ and $\mathbf{y}$ while $H(\mathbf{u} \mid \mathbf{y})$ denotes the conditional entropy of $\mathbf{u}$ given the new knowledge on the query $\mathbf{y}$. The utility entropy before the query is

$$H(\mathbf{u}) = \frac{1}{2}\log_2(|\Sigma_{\mathbf{u}}|) + \frac{\mu}{2}\log_2(2\pi e).$$

As illustrated in Figure 4(a), $H(\mathbf{u}) = I(\mathbf{u};\mathbf{y}) + H(\mathbf{u} \mid \mathbf{y})$, thus maximizing the mutual information is equivalent to minimize the conditional entropies, i.e.,

$$\text{argmax}_{\mathbf{y}}\{I(\mathbf{u};\mathbf{y})\} = \text{argmin}_{\mathbf{y}}\{H(\mathrm{u} \mid \mathbf{y})\},$$

where the postquery conditional entropy is

$$H(\mathbf{u}|\mathbf{y}) = \frac{1}{2}\log_2(|\Sigma_{\tilde{\mathbf{u}}}|) + \frac{\mu}{2}\log_2(2\pi e).$$

Let us denote a scalar $\nu_{\mathbf{u}} \equiv |\Sigma\mathbf{u}|$. By the variational principle, the gradient on the entropy (constant ignored) bears the following form:

$$\delta_{\mathbf{u}} \equiv \partial(-\log \nu_{\mathbf{u}}) = -\frac{\partial \nu_{\mathbf{u}}}{\nu_{\mathbf{u}}}, \tag{6}$$

which reflects the reduction of the entropy due to $\mathbf{y}$, i.e., the difference between $H(\mathbf{u} \mid \mathbf{y})$ and $H(\mathbf{u})$.

Theorem 1. Utility-Driven Differential Information Maximization (DIM)

The query vector $\mathbf{f}$ in (1) that maximizes the mutual information $I(\mathbf{u};\mathbf{y})$ can be derived via the following optimizer:

$$\text{argmax}_{\mathbf{f}}\,\delta_{\mathbf{u}}(\mathbf{f}) = \text{argmax}_{\mathbf{f}}\frac{\mathbf{f}^T\,\Omega\,\mathbf{f}}{\mathbf{f}^T[\Sigma_{\mathbf{x}} + \sigma^2\mathbf{I}]\mathbf{f}}, \tag{7}$$

where

$$\Omega \equiv \Sigma_{\mathbf{x}}\mathbf{U}\Sigma_{\mathbf{u}}^{-1}\mathbf{U}^T\Sigma_{\mathbf{x}} \tag{8}$$

denotes the utility amplification matrix (UAM). The optimal solution can be computed from the principal eigenvectors of the generalized eigenvalue decomposition: $\mathbf{eig}(\,\Omega, \Sigma_{\mathbf{x}} + \sigma^2\mathbf{I}\,)$.

Proof of theorem 1

We shall proceed with the proof of Theorem 1 for the scalar case, and then show how to extend the proof to the vector case.

Proof of the scalar case

Assume that $\mathbf{U} \in \Re^{M\times 1}$ and $\mathbf{u} \in \Re$, then the UAM [see (8)] has a simpler form:

$$\Omega \equiv \sigma_{\mathbf{u}}^{-2}\Sigma_{\mathbf{x}}\mathbf{U}\mathbf{U}^T\Sigma_{\mathbf{x}}. \tag{9}$$

Also, for the scalar case,

$$\sigma_{\tilde{\mathbf{u}}}^2 = \mathbf{U}^T\Sigma_{\mathbf{x}}\mathbf{U} = \nu_{\mathbf{u}} \text{ and } \sigma_{\tilde{\mathbf{u}}}^2 = \mathbf{U}^T\Sigma_{\tilde{\mathbf{x}}}\mathbf{U}. \tag{10}$$

Via (6),

$$\delta_{\mathbf{u}} = -\frac{\partial\nu_{\mathbf{u}}}{\nu_{\mathbf{u}}} = -\frac{\sigma_{\tilde{\mathbf{u}}}^2 - \sigma_{\mathbf{u}}^2}{\sigma_{\mathbf{u}}^2} = \frac{\sigma_{\mathbf{u}}^2 - \sigma_{\tilde{\mathbf{u}}}^2}{\sigma_{\mathbf{u}}^2} \tag{11}$$

and it follows that

$$\begin{aligned}\delta_{\mathbf{u}} &\underset{Eq.\,10}{=} \frac{\mathbf{U}^T(\Sigma_{\mathbf{x}} - \Sigma_{\tilde{\mathbf{x}}})\mathbf{U}}{\sigma_{\mathbf{u}}^2} \\ &\underset{Eq.\,2}{=} \frac{\mathbf{U}^T\Sigma_{\mathbf{x}}\mathbf{f}\mathbf{f}^T\Sigma_{\mathbf{x}}\mathbf{U}}{\sigma_{\mathbf{u}}^2(\mathbf{f}^T\Sigma_{\mathbf{x}}\mathbf{f}\ +\ \sigma^2)} \\ &\underset{\|f\|=1}{=} \frac{\mathbf{f}^T\Sigma_{\mathbf{x}}\mathbf{U}\mathbf{U}^T\Sigma_{\mathbf{x}}\mathbf{f}}{\sigma_{\mathbf{u}}^2(\mathbf{f}^T[\Sigma_{\mathbf{x}} + \sigma^2\mathbf{I}]\mathbf{f})} \\ &\underset{Eq.\,9}{=} \frac{\mathbf{f}^T\,\Omega\,\mathbf{f}}{\mathbf{f}^T[\Sigma_{\mathbf{x}} + \sigma^2\mathbf{I}]\mathbf{f}}.\end{aligned} \tag{12}$$

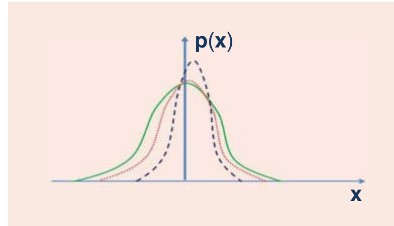


FIGURE 3. The Gaussian distribution of the original privacy function is shown by the solid green curve with the initial variance, say, $\sigma_{\tilde{p}}^2 = 30$ as exemplified in our DIP case study. Two post-query and narrower distribution curves are: a) dotted red curve with the variance reduced to $\sigma_{\tilde{\mathbf{p}}}^2 = 29.3535$ after a less intrusive query y, and b) dashed blue curve with a further reduced variance of $\sigma_{\tilde{\mathbf{p}}'}^2 = 25.7168$ after a more intrusive query y'.

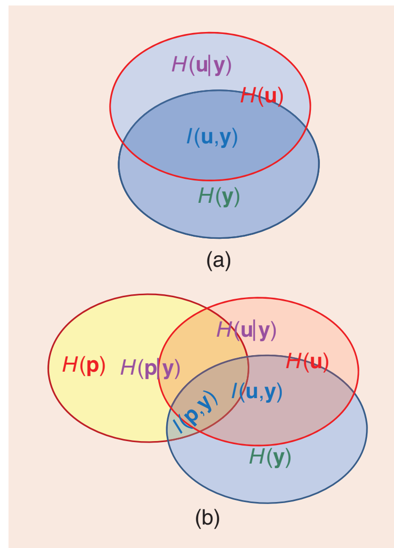


FIGURE 4. (a) From the utility's perspective, maximizing the mutual information $I(\mathbf{u};\mathbf{y})$ is the same as minimizing the conditional entropy $H(\mathbf{u} \mid \mathbf{y})$, given a query $\mathbf{y}$. (b) For the optimal utility-privacy tradeoff, we want to find a query $\mathbf{y}$ to yield larger $I(\mathbf{u};\mathbf{y})$ and smaller $I(\mathbf{p};\mathbf{y})$. This calls for an optimization metric called *DIG*:I $(\mathbf{u};\mathbf{y}) - I(\mathbf{p};\mathbf{y})$.

Proof of the vector case

We shall show that there is an orthogonalization procedure that can be used to assure that no intercomponent redundancy may exist (i.e., zero mutual information). Then the additive property that $H(\mathbf{u}) = \sum_i H(\mathbf{u}_i)$ allows each component of $\mathbf{u}$ to be treated individually, just like the scalar case. The orthogonalization hinges upon a proper transformation matrix, denoted by Γ, i.e., replacing $\mathbf{U}$ by $\mathbf{U}\Gamma$ for $\Sigma_{\mathbf{u}}$ in (4) and for Ω in (8). Note that

$$\begin{aligned}\Omega^{\text{new}} &= \Sigma_{\mathbf{x}}\mathbf{U}\Gamma[\Gamma^T\mathbf{U}^T\Sigma_{\mathbf{x}}\mathbf{U}\Gamma]^{-1}\Gamma^T\mathbf{U}^T\Sigma_{\mathbf{x}}\\ &= \Sigma_{\mathbf{x}}\mathbf{U}[\mathbf{U}^T\Sigma_{\mathbf{x}}\mathbf{U}]^{-1}\mathbf{U}^T\Sigma_{\mathbf{x}}\\ &= \Omega^{\text{old}}.\end{aligned}$$

It is important to note that Ω remains invariant with respect to any post-transformation. Thus, in our analysis, we shall pretend as if $\mathbf{U}$ were already preorthogonalized, i.e., $\mathbf{U}^T\Sigma_{\mathbf{x}}\mathbf{U} = \Lambda_{\mathbf{u}}$, where $\Lambda_{\mathbf{u}} \in \Re^{\mu\times\mu}$ is a diagonal matrix with diagonal elements: $\{\sigma^2_{\mathbf{u}_i}\}$. Let the ith column of $\mathbf{U}$ be denoted by $\mathbf{U}_i$ and then

$$\begin{aligned}\Omega &= \Sigma_{\mathbf{x}}\mathbf{U}\Lambda_{\mathbf{u}}^{-1}\mathbf{U}^T\Sigma_{\mathbf{x}}\\ &= \sum_i \Sigma_{\mathbf{x}}\mathbf{U}_i\sigma^{-2}_{\mathbf{u}_i}\mathbf{U}_i^T\Sigma_{\mathbf{x}}.\end{aligned}\quad (13)$$

The proof is now trivial having the following additive property:

$$H(\mathbf{u}) = \sum_i H(\mathbf{u}_i) = \sum_i \delta^{(i)}_{\mathbf{u}}.\quad (14)$$

Fidelity preservation query

A special situation for the aforementioned analysis is when the design objective is for the fidelity preservation. For this case, $\mathbf{U} = \mathbf{I}$ and it follows that $\Omega \equiv \Sigma_{\mathbf{x}}$ and (7) becomes

$$\text{argmax}_{\mathbf{f}}\ \delta_{\mathbf{u}}(\mathbf{f}) = \text{argmax}_{\mathbf{f}}\frac{\mathbf{f}^T\Sigma_{\mathbf{x}}\mathbf{f}}{\mathbf{f}^T[\Sigma_{\mathbf{x}}+\sigma^2\mathbf{I}]\mathbf{f}},\quad (15)$$

whose optimal solution is again the same as PCA, just like (3).

Differential utility/cost analysis (DUCA)

With reference to the Venn diagram in Figure 4(b), the basic differential information gain (DIG) can be characterized as the difference of the areas corresponding to the utility and privacy mutual information

$$\text{DIG} = I(\mathbf{u};\mathbf{y}) - I(\mathbf{p};\mathbf{y}).\quad (16)$$

Since the entropy (or information) is a log function of its corresponding covariances, thus "differential gain" in information is in some sense corresponding to the "ratio gain" in covariance.

While the default setting is $\alpha = 1$ and $\beta = 1$, it serves many practical purposes to adopt a more flexible variant:

$$\text{DIG} = \alpha I(\mathbf{u};\mathbf{y}) - \beta I(\mathbf{p};\mathbf{y})\quad (17)$$

so as to further broaden the application scope. For example,

- First, α and β are adjustable to account for the relative reward/penalty associated with the utility gain versus the privacy loss.
- Second, for constrained optimization problems such as "information bottleneck" [4] and "privacy funnel" [2], α or β play the role of being the Lagrangian multipliers.

We have previously focused on the utility-based (mutual) information optimization and, obviously, the same analysis carries through to the privacy analysis. (Detail omitted.) To study the joint utility-privacy tradeoff, we define the privacy amplification matrix (PAM) as follows:

$$\Pi = \Sigma_{\mathbf{x}}\mathbf{P}\Lambda_{\mathbf{p}}^{-1}\mathbf{P}^T\Sigma_{\mathbf{x}}.\quad (18)$$

Theorem 2. DUCA: Joint Utility-Privacy Optimization

The query vector $\mathbf{f}$ in (1) that maximizes the (weighted) "differential information gain," $\alpha I(\mathbf{u};\mathbf{y}) - \beta I(\mathbf{p};\mathbf{y})$, can be derived via the following optimizer:

$$\text{argmax}_{\mathbf{f}}\frac{\mathbf{f}^T[\alpha\Omega - \beta\Pi]\mathbf{f}}{\mathbf{f}^T[\Sigma_{\mathbf{x}}+\sigma^2\mathbf{I}]\mathbf{f}}.\quad (19)$$

The optimal solution can be computed from the principal eigenvectors of $\mathbf{eig}(\alpha\Omega - \beta\Pi, \Sigma_{\mathbf{x}} + \sigma^2\mathbf{I})$. □

Example 2. The DIP (continued)

Note that $\sigma^2_{\mathbf{u}} = 6$ and $\sigma^2_{\mathbf{p}} = 30$, via (8) and (18)

$$\Omega = \begin{bmatrix}4/6 & 8/6\\ 8/6 & 16/6\end{bmatrix} \text{ and } \Pi = \begin{bmatrix}196/30 & -224/30\\ -224/30 & 256/30\end{bmatrix}.$$

Optimal PP query

The optimal query vector is $\mathbf{f} = [0.727\ \ 0.687]^T$. This boosts the DIG from DIG = 1.61 (no query) to DIG = 2.97 (with query). Moreover, it improves the utility variance by fourfold (from six down to 1.54) while keeping the privacy loss to within 6% (from 30 down to 29.83). Pictorially, in Figure 5(a), the variance of the distribution narrows from the blue "□"s to the red "*"s.

Privacy intrusive query

From the intruder's perspective, it makes sense to set $\alpha = 0$ and $\beta = -1$ because its only objective is on the privacy attack, paying no attention to the utility gain. This results in a "most intrusive" query vector: $\mathbf{f} = [0.695\ \ -0.719]^T$, with the post-query distribution narrowing sideways (from the blue "□"s to the red "*"s), indicating that the privacy variance is severely compromised (from 30 down to 1.87) while the utility remains little affected (from six down to 5.86). □

Discriminant component analysis (DCA) machine learning and variants

The utility projection matrix $\mathbf{U}$ is sometimes known in advance but sometimes not. In machine learning, it is possible to develop learning algorithms to estimate $\mathbf{U}$ and Ω from the training data set, made available during the learning phase:

$$[\mathcal{X},\mathcal{Y}] = \{[\mathbf{x}_1,y_1],[\mathbf{x}_2,y_2],\ldots,[\mathbf{x}_N,y_N]\},$$

where the teacher values, denoted as y_i, represent two types of class labels: 1) the utility-classes, e.g., the family income level: H/M/L (say, high/middle/low), and 2) the privacy-classes, e.g., the income disparity between the couple (i.e., who earns more).

Represent UAM by between-class scatter matrices

The "center-adjusted" scatter matrix is [5]

$$\bar{\mathbf{S}} \equiv \bar{\mathbf{X}}\bar{\mathbf{X}}^T = \sum_{i=1}^{N}[\mathbf{x}_i - \vec{\mu}][\mathbf{x}_i - \vec{\mu}]^T,\quad (20)$$

which assumes the role of the covariance matrix $\Sigma_{\mathbf{x}}$ in the estimation context. The classic unsupervised learning PCA algorithm is typically computed from $\bar{\mathbf{S}}$. More exactly, the PCA subspace is represented by a projection matrix $\mathbf{W}_{\text{PCA}} \in \Re^{M\times m}$:

$$\mathbf{W}_{\text{PCA}} = \underset{\{\mathbf{W}:\mathbf{W}^T\mathbf{W}=\mathbf{I}\}}{\text{argmax}}\ \text{tr}(\mathbf{W}^T\bar{\mathbf{S}}\mathbf{W}).\quad (21)$$

The PCA solution can be computed from the m principal eigenvectors of $\mathbf{eig}(\bar{\mathbf{S}})$.

In supervised learning, the scatter matrix $\bar{\mathbf{S}}$ can be further divided into two additive parts [5]

$$\bar{\mathbf{S}} = \mathbf{S}_B + \mathbf{S}_W,$$

where the within-class scatter matrix $\mathbf{S}_W$ is defined as

$$\mathbf{S}_W = \sum_{\ell=1}^{L} \sum_{j=1}^{N_\ell} [\mathbf{x}_j^{(\ell)} - \vec{\mu}_\ell][(\mathbf{x}_j^{(\ell)} - \vec{\mu}_\ell]^T. \tag{22}$$

The between-class scatter matrix $\mathbf{S}_B$ is defined as

$$\mathbf{S}_B = \sum_{\ell=1}^{L} N_\ell\, [\vec{\mu}_\ell - \vec{\mu}][\vec{\mu}_\ell - \vec{\mu}]^T \tag{23}$$

where N_ℓ denotes the number of training vectors associated with the ℓth class, $\vec{\mu}_\ell$ denotes the centroid of the ℓth class, for $l = 1, \ldots, L$, and L denotes the number of different classes. The between-class scatter matrix represents the so-called signal subspace as it is formed from the L class-discriminating vectors learnable from the data set.

To facilitate the training of Ω from the supervised data set, we adopt two coordinate transformations to simplify the mathematical analysis.

- **Orthogonalization Transformation on U.** Without loss of generality, we assume that $\mathbf{U}$ is already pre-orthogonalized in (8), which can then be expressed as

$$\Omega = \Sigma_{\mathbf{x}} \tilde{\mathbf{U}} \tilde{\mathbf{U}}^T \Sigma_{\mathbf{x}} \tag{24}$$

where $\tilde{\mathbf{U}} \in \Re^{M \times \mu}$ with column vectors defined as $\tilde{\mathbf{U}}_i = \sigma_{\mathbf{u}_i}^{-1} \mathbf{U}_i,\ i = 1, \ldots, \mu$.

- **Whitening Transformation on x.** Let $\Sigma_{\mathbf{x}}^{\frac{1}{2}}$ denote the square-root of the covariance matrix $\Sigma_{\mathbf{x}}$ i.e., $\Sigma_{\mathbf{x}} = \Sigma_{\mathbf{x}}^{\frac{1}{2}} \Sigma_{\mathbf{x}}^{\frac{T}{2}}$. By transforming the original vector space to a "canonical" (or "whitened") vector space via

$$\mathbf{x}' = \Sigma_{\mathbf{x}}^{-\frac{1}{2}} \mathbf{x}, \tag{25}$$

the new covariance matrix becomes an identity matrix, i.e., $\Sigma_{\mathbf{x}'} = \mathbf{I}$. It follows that

$$\hat{\Omega}' = \Sigma_{\mathbf{x}'} \tilde{\mathbf{U}}' \tilde{\mathbf{U}}'^T \Sigma_{\mathbf{x}'} = \tilde{\mathbf{U}}' \tilde{\mathbf{U}}'^T, \tag{26}$$

where

$$\tilde{\mathbf{U}}' \equiv \Sigma_{\mathbf{x}}^{\frac{T}{2}} \tilde{\mathbf{U}}. \tag{27}$$

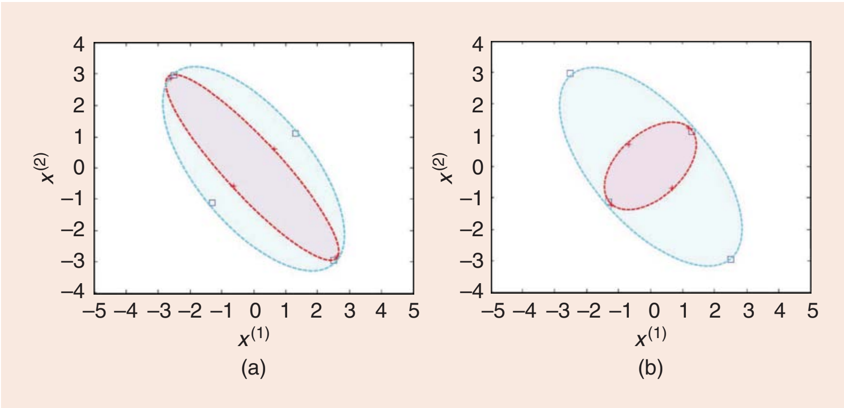


FIGURE 5. A display of DUCA-reduced covariance matrices for the DIP example, where $\mathbf{x} = [x^{(1)}\ x^{(2)}]^T$. (a) PP query, with the setting $\alpha = 1$ and $\beta = 1$, and (b) privacy-intrusive query, with $\alpha = 0$ and $\beta = -1$.

In the whitened space, the between-class utility scatter matrix $\mathbf{S}'_{B_U}$ points to the L best utility-class-discriminating vectors, spanning an $(L-1)$–dimensional subspace. Therefore, pursuant to (26), it is natural to associate each column of $\tilde{\mathbf{U}}'$ with a vector pointing from the mass center to the centroid of a utility class. This results in

$$\mathbf{S}'_{B_U} \propto \tilde{\mathbf{U}}' \tilde{\mathbf{U}}'^T \underset{Eq.\,26}{=} \hat{\Omega}'. \tag{28}$$

Note also that

$$\Omega \underset{Eq.\,24}{\propto} \Sigma_{\mathbf{x}} \tilde{\mathbf{U}} \tilde{\mathbf{U}}^T \Sigma_{\mathbf{x}} \underset{Eq.\,27}{=} \Sigma_{\mathbf{x}}^{\frac{1}{2}} \tilde{\mathbf{U}}' \tilde{\mathbf{U}}'^T \Sigma_{\mathbf{x}}^{\frac{T}{2}} \underset{Eq.\,26}{=} \Sigma_{\mathbf{x}}^{\frac{1}{2}} \Omega' \Sigma_{\mathbf{x}}^{\frac{T}{2}}. \tag{29}$$

It follows that

$$\mathbf{S}_{B_U} \underset{Eq.\,25}{=} \Sigma_{\mathbf{x}}^{\frac{1}{2}} \mathbf{S}'_{B_U} \Sigma_{\mathbf{x}}^{\frac{T}{2}} \underset{Eq.\,28}{\propto} \Sigma_{\mathbf{x}}^{\frac{1}{2}} \hat{\Omega}' \Sigma_{\mathbf{x}}^{\frac{T}{2}} \underset{Eq.\,29}{=} \hat{\Omega}. \tag{30}$$

Utility-driven DIM-DCA for supervised learning

Now we are ready to establish a machine learning variant, called DIM-DCA, corresponding to (7).

Algorithm 1. Utility-driven DCA learning algorithm

The optimization formulation of DIM-DCA involves searching for the projection matrix $\mathbf{W}_{\mathrm{DCA}} \in \Re^{M \times m}$:

$$\mathbf{W}_{\mathrm{DCA}} = \underset{\{\mathbf{W}:\mathbf{W}^T[\bar{\mathbf{S}}+\rho\mathbf{I}]\mathbf{W}=\mathbf{I}\}}{\operatorname{argmax}} \operatorname{tr}(\mathbf{W}^T \mathbf{S}_{B_U} \mathbf{W}).$$

(Note that ρ here assumes the role of the variance σ^2.) The optimal DCA solution can be derived from the principal eigen-subspace of the following "discriminant matrix" [6]:

$$[\bar{\mathbf{S}} + \rho \mathbf{I}]^{-1} \mathbf{S}_{B_U}. \tag{31}$$

Equivalently, they can be derived from the first m principal eigenvectors of

$$\mathbf{eig}(\ \mathbf{S}_{B_U},\ \bar{\mathbf{S}} + \rho \mathbf{I}). \tag{32}$$

The extracted queries are rank-ordered according to their "signal to power ratios," which are equivalent to their corresponding eigenvalues:

$$\lambda_i = \begin{cases} \dfrac{\mathbf{v}_i^T \mathbf{S}_{B_U} \mathbf{v}_i}{\mathbf{v}_i^T\ (\bar{\mathbf{S}} + \rho \mathbf{I})\ \mathbf{v}_i} & \text{if} \quad i = 1, \ldots, L-1 \\ 0 & \text{if} \quad i \geq L. \end{cases} \tag{33}$$

□

Metric for interclass separability

The trace-norm of the discriminant matrix, defined in (31), may be used as a simple metric to measure the inter-class separability of a supervised data set. It offers a convenient tool to evaluate the the suitability of a certain similarity function (or kernel function [3]) to be chosen for nonlinear data analysis (see the next section).

Theoretical connection between two variants of DCA

In [6], another variant of DCA was developed for finding the optimal subspace projection matrix via the principal eigenvectors of

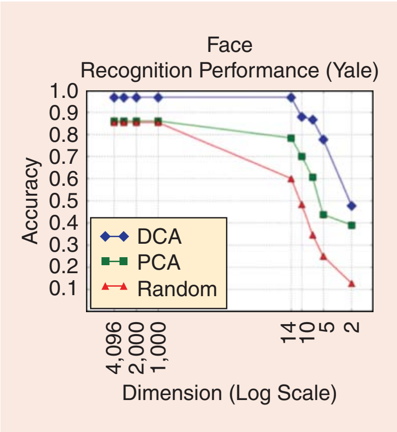


FIGURE 6. The FR accuracies on the Yale data set, with respect to different reduced-dimensions via DCA, PCA, and random projection. (Figure courtesy of T. Chanyaswad.)

$$\mathbf{eig}(\bar{\mathbf{S}}, \mathbf{S}_{W_U} + \rho \mathbf{I}), \tag{34}$$

where $\mathbf{S}_{W_U}$ denotes the within-(utility)-class scatter matrix. This variant is an extension of Fisher's LDA and is indeed a very close sibling of the DIM-type DCA. To prove that the Fisher-type DCA is exactly the same as the DIM-type DCAs, when $\rho = 0$, we let $\{\lambda_i, \mathbf{v}_i\}$ and $\{\lambda_i', \mathbf{v}_i'\}$ respectively denote the eigenvalues/eigenvectors for DIM-DCA [see (32)] and Fisher-DCA [see (34)]. It can then be shown that $\mathbf{v}_i' = \mathbf{v}_i$ and $\lambda_i' = (1 - \lambda_i)^{-1}$. The latter guarantees $\{\lambda_i'\}$ and $\{\lambda_i\}$ to be sorted in the same order, thus verifying the equivalence.

PPFR simulation results

Apply PCA and DCA to the Yale data set for PP face recognition (FR) (PPFR) applications, we have the following observations:

- **PCA/DCA Classification Accuracies.** There are only $L-1$ meaningful eigenvectors, because **rank** $(\mathbf{S}_{B_U}) \le L-1$. Note also that usually $L \ll M$, it implies that the DCA eigen-components can enjoy a win-win advantage in improving privacy without sacrificing utility.
  - First, the $L-1$ principal components can capture key features fully adequate for very high performance, as evidenced by the performance curves shown in Figure 6. Note that DCA far outperforms PCA and random projection in terms of FR accuracies.
  - The DCA dimension reduction results in removal of a large proportion of components, making it an effective compression tool for privacy preservation.
- **Data Visualization by PCA/DCA Projection.** The high-dimensional Yale data set may be visualized by means of two-dimensional PCA or DCA subspace projection. Figure 7(a) displays the PCA visualization, showing that many classes are nonseparable by PCA. In contrast, the DCA visualization in Figure 7(b) shows very well separated classes. In fact, many data points from the same class align almost perfectly, sometimes making the whole class of data projected to a single point.

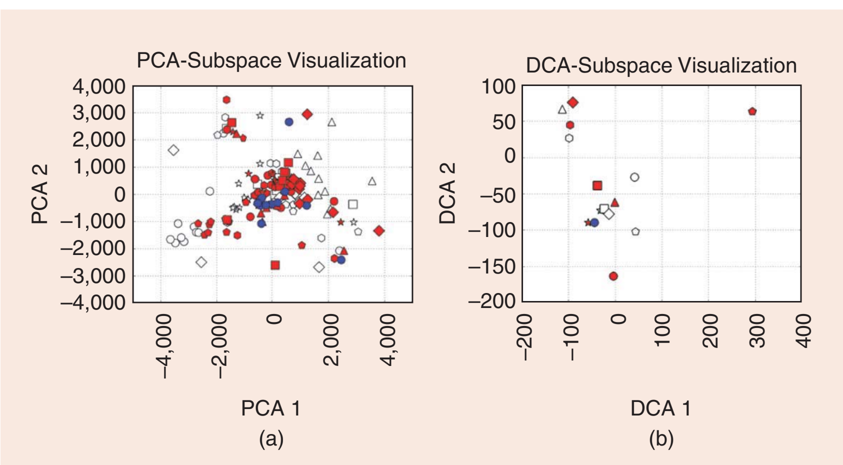


FIGURE 7. The high-dimensional Yale data set may be visualized by means of two-dimensional (a) PCA or (b) DCA subspace projection. Each mark represents a data point, and data points in the same class share the same shape and color. (Figure courtesy of T. Chanyaswad.)

Privacy-driven desensitized DCA via ridge DCA (RDCA)

In the previous section, the utility-driven learning algorithms are good for scenarios where the intended utility is well defined but the privacy policy is still open. Conversely, there are other scenarios where the intended utility is yet to be determined but the privacy policy is already pre-defined. This calls for a DCA variant tailored designed for the extraction of desensitized components. To this end, we further incorporate another ridge parameter ρ' to regulate the (privacy) signal matrix $\mathbf{S}_{B_P}$, resulting in the following privacy-driven learning algorithm [7].

Algorithm 2: Privacy-driven ridge DCA algorithm

Find the projection matrix $\mathbf{W}_{\text{RDCA}} \in \Re^{M \times m}$:

$$\mathbf{W}_{\text{RDCA}} = \underset{\{\mathbf{W}: \mathbf{W}^T[\bar{\mathbf{S}} + \rho \mathbf{I}]\mathbf{W} = \mathbf{I}\}}{\text{argmax}} \text{tr}(\mathbf{W}^T[\mathbf{S}_{B_P} - \rho' \mathbf{I}]\mathbf{W}).$$

where $\mathbf{S}_{B_P}$ denotes the between-(privacy)-class scatter matrix and ρ' is a small positive value. The optimal RDCA solution can be derived from the m eigenvectors corresponding to the Lth, $\ldots, (L+m-1)^{th}$ eigenvalues of

$$\mathbf{eig}(\mathbf{S}_{B_P} - \rho' \mathbf{I}, \bar{\mathbf{S}} + \rho \mathbf{I}). \tag{35}$$

The component powers are closely related to their corresponding eigenvalues:

$$P(\mathbf{v}_i) \approx -\frac{\rho'}{\lambda_i} - \rho, \text{ for } i \ge L. \tag{36}$$

□

Equation (36) assures that the desensitized components can be orderly extracted according to their eigenvalues, just like PCA. This is why RDCA is sometimes referred to as *desensitized DCA* or, more simply, *desensitized PCA*. Let us highlight some key properties of RDCA's eigen-components:

- *Signal-Subspace Components,* i.e., when $i < L$: The first $L-1$ eigen-components are potentially most

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

intrusive, so they must be filtered out for the privacy sake.

- *Noise Subspace Components,* i.e., when $i \geq L$: In contrast, $\mathbf{v}_i^T \mathbf{S}_{B_P} \mathbf{v}_i \simeq 0$ for all $i \geq L$, indicating that they do not leak sensitive information. Moreover, their utilizable component powers are rank ordered by the corresponding eigenvalues λ_i.

Antirecognition utility maximization results

Antirecognition utility maximization (ARUM) is an exemplifying application scenario, in which the privacy policy calls for deidentification. More exactly, while we want to conceal the person's identity, we would like to retain as much information as possible to serve other purposes such as 1) eyeglasses detection or 2) mood detection, e.g., happy versus sad faces. Further ARUM applications may include, for example, anomaly/intruder detection without leaking the exact identity of the user.

As an example, we have applied DCA and RDCA to the Yale data set, with 15 different persons (i.e., $L = 15$), and compared their performances for both PPFR and ARUM applications.

- **Eigenfaces in the Signal-Subspaces of DCA and RDCA.** Note that, when ρ' is relatively small, DCA and RDCA share almost the same signal subspace. Specifically, their 14 $(= L - 1)$ principal eigenfaces are very similar. How to treat the 14 eigenfaces depends on the intended utility/privacy goals:
  – For PPFR, the principal eigenfaces alone are sufficient to yield high accuracies shown in Figure 6.
  – For ARUM, on the other hand, the same principal eigenfaces are deemed to be most intrusive and, therefore, they should be cast away for the sake of desensitization.
- **Eigenfaces in the Noise-Subspace.** Shown in Figure 8(a) are the first five desensitized DCA eigenfaces and, in Figure 8(b), the first five RDCA eigenfaces. Each of the noise-subspace components yields a very low classification accuracy around 6.6%, no better than random guess out of $L = 15$ choices. This confirms that each of them is perfectly desensitized and contains no useful information for authentication.

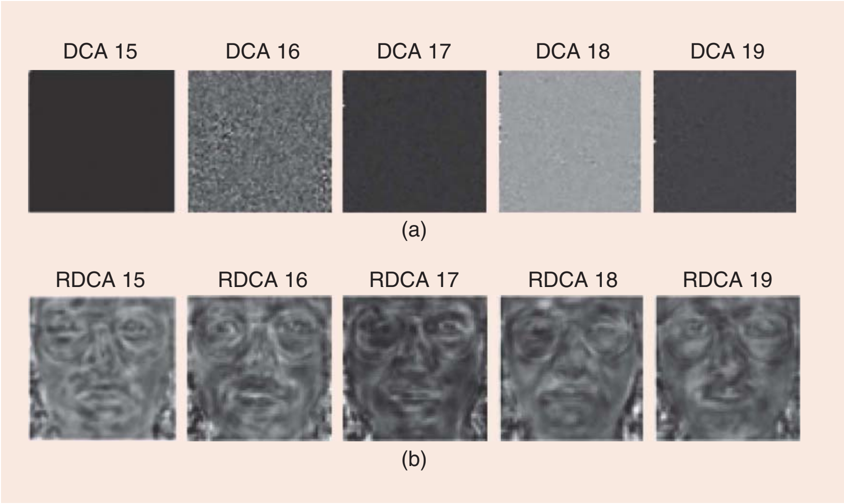


FIGURE 8. The DCA and RDCA basically have the same principal (14) eigen-faces in their signal subspace (not shown here). However, they have completely different eigen-faces in the noise subspace. (a) The next five privatized DCA-eignfaces are random and useless. (b) In contrast, rich information is revealed by the five highest-power desensitized RDCA-eigenfaces, corresponding to the 15th–19th eigenvalues (r'= 0.05). (Figure courtesy of T. Chanyaswad and A. Filipowicz.)

Reconstructed images by DCA and RDCA

The performance distinction lies in the power-based rank-ordering (or lack of it) of the DCA's and RDCA's desensitized components:

- Figure 8(a) shows the DCA's desensitized eigenfaces, which apparently contain no useful information.
- Figure 8(b) shows the RDCA's desensitized eigenfaces, which exhibit much higher component powers, rendering them possibly amenable to ARUM-type applications.

By the following ARUM-type example, we show how to harness the reconstruction for some useful applications.

- Shown in Figure 9(a) is the original image of the first sample in the Yale data set, which is sampled/represented by a full-dimensional vector $(M = 4{,}096)$.
- Figure 9(b) displays the reconstructed face images using 3,986 DCA components.
- Figure 9(c) displays the reconstructed face images via 3,986 (power-sorted) RDCA components.

By comparing the two reconstructed images with the original image, we note that the RDCA outperforms DCA for both the eyeglasses detection and mood detections. (More visibly so for the former but somewhat subjective for the latter.)

Quantitative tradeoff analysis via RDCA

An experimental study on an in-house glasses data set (with seven different persons) was conducted for an ARUM-type application[7], where 1) the utility is to determine whether or not a person wears glasses and 2) the privacy involves concealing the person's identity, i.e., deidentification.

Our study involved 1,000 trials. RDCA appears to be promising: Upon the RDCA desensitization, the privacy accuracy drops significantly from 97.6% to 44.4%, implying a much improved protection, while its utility accuracy remains fairly high, from 98.3% to 95.5%.

DUCA machine learning and variants

It would be idealistic if we could benefit from having both types of the teacher values made available during the learning phase, one for the utility class and one for privacy class. Together with the training data set, these teacher values can be used to produce two between-class scatter matrices, $\mathbf{S}_{B_U}$ (for utility) and $\mathbf{S}_{B_P}$ (for privacy), which can, in turn, be used to estimate Ω and Π, respectively.

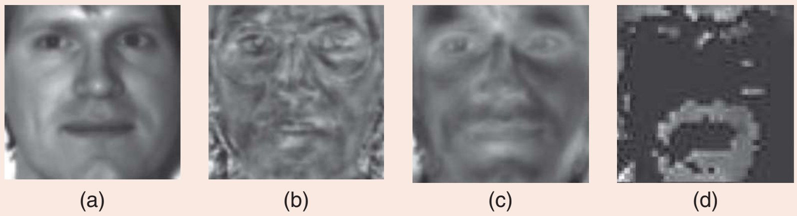


FIGURE 9. The original and reconstructed image of the first sample of the Yale data set. (a) The original face image, (b) the DCA reconstructed image, (c) the RDCA reconstructed image (with $\rho' = 0.05$), and (d) the reconstructed image from 399 selected wavelet components. (Figure courtesy of T. Chanyaswad and A. Filipowicz.)

DUCA: Joint utility-privacy learning algorithm

For utility-privacy tradeoff analysis, a natural combination of Algorithm 1 and (19) leads to the following DUCA supervised learning algorithm.

Algorithm 3. DUCA supervised learning algorithm

Find the projection matrix $\mathbf{W}_{\text{DUCA}} \in \Re^{M \times m}$ such that

$$\mathbf{W}_{\text{DUCA}} = \underset{\{\mathbf{W}:\mathbf{W}^T[\bar{\mathbf{S}}+\rho\mathbf{I}]\mathbf{W}=\mathbf{I}\}}{\operatorname{argmax}} \operatorname{tr}\left(\mathbf{W}^T[\alpha\mathbf{S}_{B_U} - \beta\mathbf{S}_{B_P}]\mathbf{W}\right).$$

whose solution can be derived from the m principal eigenvectors of

$$\mathbf{eig}\left(\alpha\mathbf{S}_{B_U} - \beta\mathbf{S}_{B_P}, \bar{\mathbf{S}}+\rho\mathbf{I}\right). \quad (37)$$

□

Note that there are only $L+C-2$ meaningful eigenvectors, because $\mathbf{rank}\left(\alpha\mathbf{S}_{B_U} - \beta\mathbf{S}_{B_P}\right) \le L+C-2$ where L and C denote the numbers of utility and privacy classes, respectively.

Example 3

Machine learning for DIP. Let us revisit the DIP example, now with two more features. Suppose that we are given a training data set:

$$\{\mathcal{X}\} = \begin{bmatrix}11\\7\\1\\2\end{bmatrix}\begin{bmatrix}18\\8\\2\\-1\end{bmatrix}\begin{bmatrix}17\\5\\-1\\-1\end{bmatrix}\begin{bmatrix}4\\10\\-1\\-4\end{bmatrix}\begin{bmatrix}5\\6\\2\\2\end{bmatrix}\begin{bmatrix}4\\7\\1\\-1\end{bmatrix}\begin{bmatrix}1\\2\\-1\\1\end{bmatrix}\begin{bmatrix}4\\1\\1\\-1\end{bmatrix},$$

with the utility/privacy teacher labels, denoted by

$$\{\mathcal{Y}\} = \begin{bmatrix}H\\\triangle\end{bmatrix}\begin{bmatrix}H\\\triangle\end{bmatrix}\begin{bmatrix}M\\\triangle\end{bmatrix}\begin{bmatrix}M\\\nabla\end{bmatrix}\begin{bmatrix}M\\\nabla\end{bmatrix}\begin{bmatrix}M\\\nabla\end{bmatrix}\begin{bmatrix}L\\\nabla\end{bmatrix}\begin{bmatrix}L\\\triangle\end{bmatrix}$$

where "H/M/L" denotes the three (high/middle/low) utility classes (i.e., family income) and "△/▽" denotes the two privacy classes (i.e., who earns more between the couple).

Recall that, via (20), the scatter matrix can be learned from the given data set $\{\mathcal{X}\}$ and UAM and CAM can be learned, via (23), by further incorporating their respective class labels. Let us set the query vector as $\mathbf{f} = [.9\ .4\ .1\ -.1]^T$, whose projection to the two-dimensional subspace is marked as "♡." Given the full-dimensional query it is clear that it belongs to the middle (utility) class, because $\mathbf{u}(\mathbf{x}) = x_1 + x_2 = 15$ and to the "△" privacy class, because $\mathbf{p}(\mathbf{x}) = x_1 - x_2 = 5 > 0$. The objective of CP design is to find a dimension-reduced query to correctly classify the utility class but not to reveal its privacy class. By comparing the PCA, DCA, and DUCA subspace projections, depicted in Figure 10, DUCA is noticeably the far best in meeting both objectives. □

It is worth elaborating further on the performance comparison between PCA, DCA, and DUCA.

- The two PCA principle eigenvectors of (21) are $\mathbf{f}_1 = [.984\ .174\ .039 - .016]^T$ and $\mathbf{f}_2 = [.163 - .899\ .042\ .405]^T$. As shown by the PCA projections in Figure 10(a), the utility class of the query "♡" is somewhat undetermined because it is possible to be either middle-income or high-income. More seriously, Figure 10(b) strongly hints its "△" class, potentially exposing the privacy.
- The two DCA principle eigenvectors of (32) are $\mathbf{f}_1 = [.204\ .838\ .245\ .443]^T$ and $\mathbf{f}_2 = [.221 - .535\ .733\ .357]^T$, leading to the DCA projections in Figure 10(c) and (d). The good news is that the query "♡" is correctly classified as middle-income. Moreover, when compared with PCA, cf. Figure 10(b) and (d), the two privacy classes seem to overlap more, which suggests an enhanced privacy. Indeed, although the query "♡" is leaning towards the "△" class, its uncertainty represents a noticeable improvement over PCA.
- With $\alpha = \beta = 1$, the two DUCA principle eigenvectors of (37) are: $\mathbf{f}_1 = [.142\ .872\ .168\ .438]^T$ and $\mathbf{f}_2 = [-.002\ -.114\ .682\ .723]^T$. The DUCA subspace projection enjoys the best of the two worlds:
  - As shown in Figure 10(e), the query "♡" can be confidently classified into the middle-income class.
  - As shown in Figure 10(f), the two privacy classes now overlap much more than before. Indeed, the privacy label for the query "♡" appears to be totally undecided.

Supervised DUCA-based filtering for feature selection

An alternative approach to the dimension reduction of the query vector is the feature selection strategy. Its design objective is to retain only a small number of selective features for CP. In this case, the feature vector is restricted to an indicator-type vector: $\mathbf{f} = [0 \cdots\ 0\ 1\ 0\ \cdots 0]$, where only the ith entry is nonzero. This brings about the following DUCA-scores useful for ranking the features in feature selection:

$$\text{DUCA}(i) = \frac{\alpha\mathbf{S}_{B_U}(i,i) - \beta\mathbf{S}_{B_P}(i,i)}{\bar{\mathbf{S}}(i,i)+\rho}. \quad (38)$$

Example 4

DUCA Ranking for DIP Feature Selection: According to (38), the DUCA-scores are $\text{DUCA}(1,2,3,4) = [-0.05\ 0.73\ 0.19\ \ 0.12]$. This is consistent with the previous finding that (1) x_1, being most DUCA-costly, is given the lowest weight in both $\mathbf{f}_1$ and $\mathbf{f}_2$; and (2) x_2, being most DUCA-rewarding, receives the highest weight in $\mathbf{f}_1$.

□

IEEE SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page | Qmags THE WORLD'S NEWSSTAND®

Example 5

DUCA Filtering for PPFR Application: By applying SVM to the full-dimensional Yale data set, it yields a recognition accuracy of 82%. We have also applied a utility-driven DUCA score, i.e., $-\mathbf{S}_{B_U}(i,i)/(\bar{\mathbf{S}}(i,i)+\rho)$, to select the best 399 (of 4,096) Wavelet-transformed components. We have found that the DUCA-filtered CP method actually offers a higher accuracy at 82.3%, again via SVM. At the same time, it also totally obfuscates the face images, as exemplified by Figure 9(d). In short, the DUCA-filtering feature selection is promising for PPFR since it offers PP compression without compromising the FR accuracy. □

Extension to kernel DCA and kernel DUCA

In the kernel learning models [3], $\mathbf{u}(\mathbf{x})$ and $\mathbf{p}(\mathbf{x})$ will be nonlinear functions in general. As such, it can induce an expanded solution space and thus further improve the performance. It involves a simple kernelization procedure to extend from DCA to kernel DCA [6]. For example, the discriminant matrix in (31) can be extended to the following kernel-DCA discriminant matrix:

$$[\bar{\mathbf{K}}^2+\rho\bar{\mathbf{K}}]^{-1}\mathbf{K}_{B_U} \quad (39)$$

where $\bar{\mathbf{K}}$ and $\mathbf{K}_{B_U}$ denote the kernelized counterparts of $\bar{\mathbf{S}}$ and $\mathbf{S}_{B_U}$, respectively. Again, applying eigen-space analysis to this kernelized matrix will lead to the optimal query solution in the kernel vector space.

For applications to CP problems, a kernel-DUCA discriminant matrix may be derived by substituting $\mathbf{K}_{B_U}$ by $\mathbf{K}_{B_U}-\mathbf{K}_{B_P}$ in (39). (Again, the principle eigen-subspace analysis would yield the optimal queries.) However, for such applications, the reduced dimension must be strictly lower than the original dimension M, since the CP encoding scheme must necessarily be lossy.

Recently, there has been growing interest in multikernel research[3], where a multikernel function is expressed as linear combination of many kernels: $K(\mathbf{x},\mathbf{y})=\sum_l^L \gamma_l K_l(\mathbf{x},\mathbf{y})$. In this case, the trace-norm of kernel-DCA discriminant matrix in (39) may be used as an effective evaluation criterion for finding the optimal coefficients: $\{\gamma_l,\ \ l=1,\ldots,L\}$. (Detail omitted.)

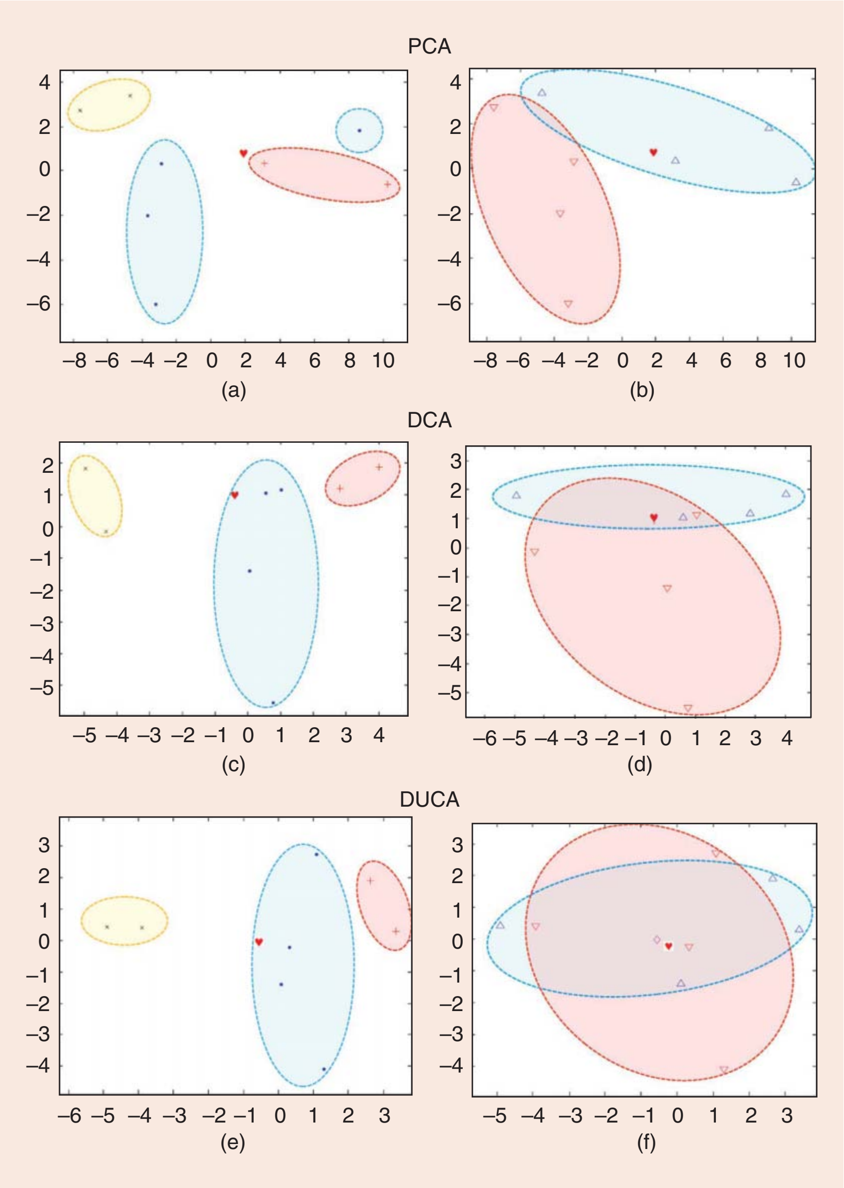


FIGURE 10. Visualization of a query, marked as ♡, mapped to the (a) and (b) optimal two-dimensional PCA, (c) and (d), DCA, and (e) and (f) DUCA subspaces. The high/middle/low utility labels are marked by $+/*/\times$, and the two privacy labels are marked by $\triangle/\nabla$. The results suggest that DUCA-subspace offers a promising approach to optimal utility-privacy tradeoff.

Tailor designed noise for privacy preservation

Pursuant to (23), the privacy matrix $\mathbf{S}_{B_P}$ may be learned from labeled training data and the original data $\vec{\mathbf{x}}$ be purposefully perturbed by noise parallel to the privacy subspace, leading to a new query: $\vec{\mathbf{y}}=\vec{\mathbf{f}}^T(\vec{x}+\epsilon)$, with colored noise covariance $\sum_\epsilon=\rho\mathbf{S}_{B_P}$. This compels the eigen-solution (37) to be modified as:

$$\mathbf{eig}(\alpha\mathbf{S}_{B_U}-\beta\mathbf{S}_{B_P},\bar{\mathbf{S}}+\rho\mathbf{S}_{B_P})$$

Intuitively speaking, such a design aims to dampen Eve's ability to intrude privacy while leaving the utility gain for Bob relatively unaffected.

(continued on page 112)

APPLICATIONS CORNER

Muhammad Salman Khan, James Stewart Jenkins,
and Nestor Becerra Yoma

Discovering New Worlds

A review of signal processing methods for detecting exoplanets from astronomical radial velocity data

Exoplanets, short for *extra solar planets*, are planets outside our solar system. They are objects with masses fewer than around 15 Jupiter-masses that orbit stars other than the sun. They are small enough so they cannot burn deuterium in their cores, yet large enough that they are not so-called dwarf planets like Pluto.

To discover life elsewhere in the universe, particularly outside our own solar system, a good starting point would be to search for planets orbiting nearby sun-like stars, since the only example of life we know of thrives on a planet we call Earth that orbits a G-type dwarf star. Furthermore, understanding the population of exoplanetary systems in the nearby solar neighborhood allows us to understand the mechanisms that built our own solar system and gave rise to the conditions necessary for our tree of life to flourish.

Signal processing is an integral part of exoplanet detection. From improving the signal-to-noise ratio of the observed data to applying advanced statistical signal processing methods (among others), to detect signals (potential planets) in the data, astronomers have tended, and continue to tend, toward signal processing in their quest of finding Earth-like planets. Methods that have been used to detect exoplanets are listed in "Discovering Exoplanets."

Digital Object Identifier 10.1109/MSP.2016.2617293
Date of publication: 11 January 2017

Discovering Exoplanets

Radial velocities: where the gravitational tug of the planet on the star is measured by analyzing the stellar spectral fingerprint to search for the Doppler shift of these lines as the star and planet orbit their common center of mass.

Transits: where the planet passes in front of the star toward our line of sight, blocking the star's light as it does so, and inducing a slight dimming of the light profile.

Transit timing variations: where the time of center of transit is measured over many transits, and variations in that time that are due to the gravitational interaction from another planet in the system can be measured.

Photometric variations: a series of methods that model variations in a star's photometric light curve to infer the presence of planets orbiting the star (e.g., planetary reflected light or thermal emission, Doppler boosting, and ellipsoidal variations).

Gravitational microlensing: where a foreground star passes across the line of sight of a far-off star, and the gravitational field of the foreground star acts as a lens to intensify the light of the background star, and also intensifies the light from a planet orbiting that star.

Pulsar timing: where the precise arrival time of the pulses of light are measured, and small differences in the timing can be introduced due to small planets orbiting these dead stars.

Direct imaging: where we point large telescopes at stars and directly image any orbiting planets.

Astrometric wobble: where we measure the position of a star on the sky and search for changes in that position, or a wobble, due to the gravitational tug of orbiting planets.

In this article, we focus on the radial velocity method of exoplanet detection, the most successful method for discovering planets orbiting the nearest stars to the sun [1]–[3]. We address basic questions such as

- How is the radial velocity data obtained?
- Why is the data nonuniformly sampled?
- What are the different signal processing methods that astronomers have been using for detecting exoplanets and what are their pros and cons?
- What is the statistical significance of signal detection?
- What are the potential directions for future research?

1053-5888/17©2017IEEE

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

Introduction

The radial velocity method works by breaking a star's light up into its constituent colors using a high-resolution echelle spectrograph. The observed stellar spectral lines can then be used as markers for the star's velocity. If the star's velocity changes, the Doppler effect tells us that electromagnetic waves are affected by this movement by presenting a shift in frequency, depicted in Figure 1. We can measure that frequency and then correct for any additional velocity shifts from noise sources, such as temperature and pressure variations in the laboratory, mechanical instabilities throughout the optical train, observational airmass chromatic effects, stellar magnetic activity affects, and stellar convective blueshift, and measure the star's radial velocity toward or away from us. Over the course of a planet's orbital period, we can measure the star's spectral redshift and blueshift, and by analyzing the amplitude, phase, shape, and period of this signal, we can understand characteristics about the companion that is causing the star's velocity variations.

Exoplanets are difficult to detect due to their extreme contrasting characteristics compared to the stars they orbit. Planets are so much smaller and fainter than their host stars that all of the aforementioned methods have a difficult time detecting them. For the radial velocity technique, which is the focus of this article, the much smaller mass of the planet compared to the star means we need spectrographs that are stable at the meters/second level to detect even the most massive planets, and to detect Earths in Earth-like orbits around sun-like stars, we need spectrographs that are stable at the centimeters/second level. This is very difficult to accomplish since small pressure and temperature variations, illumination problems, mechanical stability problems, and even the stars themselves can introduce noise in the measurements at levels higher than this.

The radial velocity data is obtained by observing a star with a telescope and feeding the light to an echelle spectrograph, as mentioned previously. We can observe the star as many times as we want in a single night and as many times as we can when the star is in the sky throughout the year. The more observations we get, the better we sample the signal of the star's radial velocity. The data is first reduced, which is a way of using calibrations to prepare the spectra for analysis, and then we can measure the velocity. The typical reduction procedure for such data is to perform a bias correction to the image, so-called debiasing, then we correct for the pixel-to-pixel variations through a process called *flatfielding*, any scattered-light is then removed from the high signal-to-noise ratio data, and the spectra can be extracted, or collapsed, into its two-dimensional format. Finally, all cosmic rays or bad pixels can be cleaned from the extracted spectrum and a highly precise wavelength correction is applied. Interested readers can refer to, e.g., [4] and [5] for more details of this process.

The radial velocity data is unevenly sampled because we can only observe the star when it is visible in the night sky. Stars are not visible all year round, as sometimes they are in the same area of sky as the sun. Furthermore, we must compete for telescope time, so we cannot always observe when we want—we are at the mercy of schedulers and proposal reviewers. Also, the number of stars we can observe per night is limited, e.g., 30–40 or more, so we cannot observe all of them every night that we actually have telescope time.

In the following sections, we review the different signal processing methods used by the astronomical community for detecting exoplanets based on radial velocity data. These methods include the Lomb–Scargle (LS) periodogram, Keplerian periodogram, prewhitening method, maximum-likelihood (ML) periodograms, Bayesian analysis, and the minimum mean square error (MMSE)-based method. All of the methods assume that $x(t)$ is the radial velocity data, t is the timestamp of observations,

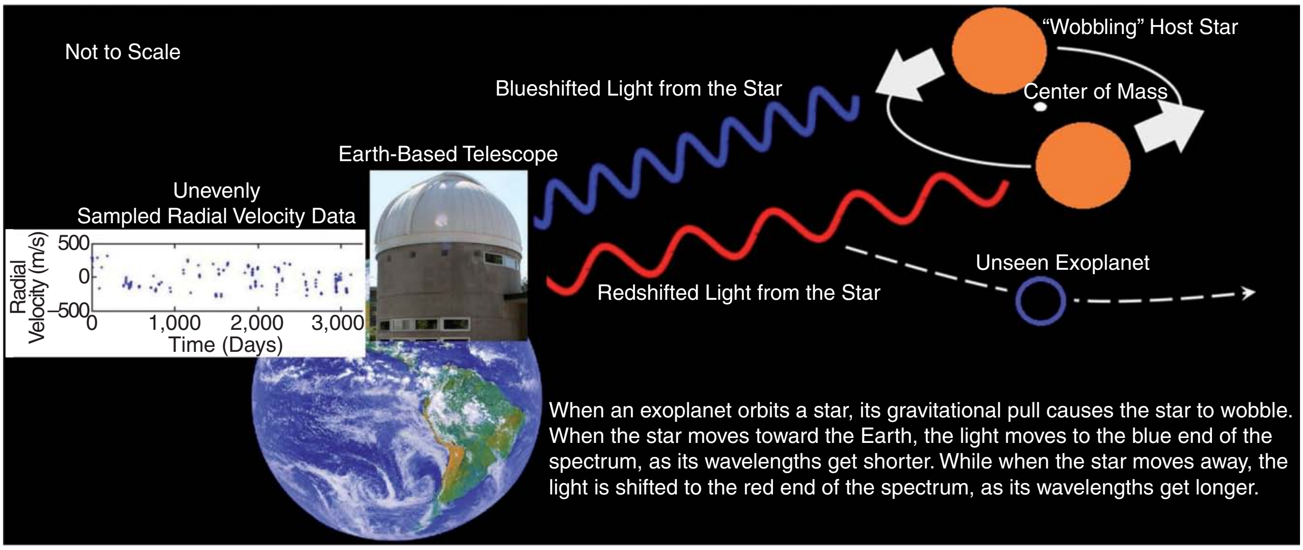


FIGURE 1. Detecting exoplanets using the radial velocity method.

i.e., $t = 1, 2, 3, \ldots, T$, where T is the total number of observations.

LS periodogram

The LS method [6] of signal detection has been extensively used in the search for exoplanets, particularly using the radial velocity technique of planet detection. In its simplest form, the method works in a Fourier-like manner by applying a number of sines and cosines to the radial velocity data $x(t)$ across a grid of frequencies chosen by the user, and the amplitude of these functions are minimized to fit the data and a power is calculated. When one of the functions provides a good match to the radial velocity time series, the power will be maximized at the selected frequency, indicating to the user that there is a signal at that frequency, and this can be visually viewed by a periodogram:

$$P_x(\omega) = \frac{1}{2}\left\{\frac{\left[\sum_{t=1}^{T} x(t)\cos\omega(t-\tau)\right]^2}{\sum_{t=1}^{T}\cos^2\omega(t-\tau)} + \frac{\left[\sum_{t=1}^{T} x(t)\sin\omega(t-\tau)\right]^2}{\sum_{t=1}^{T}\sin^2\omega(t-\tau)}\right\}, \tag{1}$$

where P_x is the periodogram powers as a function of frequency, and τ is defined as

$$\tan(2\omega\tau) = \frac{\left(\sum_{t=1}^{T}\sin 2\omega t\right)}{\left(\sum_{t=1}^{T}\cos 2\omega t\right)}.$$

Although the technique is easy to implement and fast to apply even to large data sets, it has some major drawbacks when searching for Doppler signals induced on stars from orbiting planets. For instance, not all exoplanets are found to be on circular orbits around their stars. In fact, there is a high fraction that have significant eccentricities, and once the eccentricity of the orbit is larger than ~0.6, the LS method finds it more difficult to detect these signals. Another issue with this method is that it makes the assumption there is only one signal in the data, each time the method is applied. Yet many planetary systems are found to contain more than one planet, which means the radial velocity time series should exhibit more than one signal. This assumption also underlies the Fourier transform analysis [7]. Therefore, signals must be subtracted out of the data fits, before reapplication of the method on the residuals is performed, and since the signals are generally not orthonormal, this gradient-based approach introduces problems for detecting low-mass and multiple planet systems. It is worth mentioning that in [8] a date-compensated discrete Fourier transform was proposed that gives better estimates of the power spectrum of nonuniformly sampled data, aiding in accurate determination of spectral peak heights. Finally, the LS method only considers white noise, which can be problematic since starlight often involves correlated noise.

Keplerian periodogram

Given some of the problems mentioned with the LS periodogram method of signal detection—particularly the fact that signals are not always well described by sines or cosines—the Keplerian periodogram was developed [9], [10]. The Keplerian periodogram allows the user to consider factors of noncircularity as part of the analysis when calculating the powers for the periodogram, since the chi-squared comparison used is open to any model that can be fit to the data, for instance

$$p_{Kep}(\omega) = \frac{\chi_0^2 - \chi_{Kep}^2(\omega)}{\chi_0^2}.$$

Here $p_{Kep}(\omega)$ is the power, χ_0^2 is the chi-squared for the weighted mean, and $\chi_{Kep}^2(\omega)$ is the chi-squared of the Keplerian model. The Keplerian model in this case can be written as

$$x(t) = \gamma + K[e\cos\varpi + \cos(\nu(t) + \varpi)], \tag{2}$$

where γ is the systemic offset of the data, K is the amplitude of the signal, e is the eccentricity of the orbit, ϖ is the longitude of periastron of the orbit, and $\nu(t)$ is the true anomaly of the orbit. Keplerian signals can be detected in radial velocity data following this approach, but we remind the reader that the Keplerian periodogram is open to including other models to calculate powers. In contrast to the LS method, the Keplerian method is relatively slow and complicated to apply to long time series data, but since it is more robust in detecting signals that deviate from sinusoids, it is more applicable in the search for exoplanetary systems orbiting nearby stars [11]. However, it again makes the assumption that there is only one signal in the data, which means it suffers from the same problems as the LS method if the data contains more than one signal, and there is no correlated noise component.

Prewhitening method

The method of prewhitening to search for Doppler signals in radial velocity time series is similar to the LS method, except that the method is applied in the Fourier domain to search for any signals in the data, but again using sines and cosines (e.g. [12]). As the name suggests, the method works by whitening the data as much as possible to remove all noise sources with fitted functions, until a real Doppler signal is found. The data is translated into Fourier space and a search for frequencies that pass a significance threshold is performed. The strongest signal is fit, the corresponding residual to the fit calculated, and then the process is repeated again. This goes on until the residual data is just the noise-floor of the observations, meaning no peaks are found above the significance threshold. Similar to the LS and generalized LS (GLS) methods, this is quick and easy to apply, but it has the same problems as these other two methods. However, the prewhitening part is done to clean noise from the time series, but that requires knowledge of the noise source, like aliases of real signals, or in the case of stars, stellar activity signals/timescales, and again, this is a gradient-based approach that does not consider correlated noise.

ML periodograms

Given that the aforementioned approaches focus on searching for one signal at a time in the time series when searching for planets, and none deal with correlated noise, ML periodogram methods have been developed to circumvent these issues [13], [14]. The ML periodogram method does not generate a periodogram that shows power on the y-axis, but instead it shows the log-likelihood of the model that is compared to the data at each step. In this way, any model can be compared to the data directly across a grid of frequencies or orbital periods, and the log-likelihood can be calculated for each, with a detected signal having the maximum likelihood

$$L(m|\theta) = \prod_{f=1}^{N} \frac{1}{\sqrt{2\pi(\sigma_f^2 + \sigma_I^2)}} \exp\left\{\frac{-[m_f - \nu_l(t_f)]^2}{2(\sigma_f^2 + \sigma_I^2)}\right\}. \quad (3)$$

The likelihood to be maximized can be described as in (3) with $L(m|\theta)$ being the likelihood of the data m given the model parameters θ. σ_f and σ_I represent the stellar and instrumental white noise components, respectively, and $\nu_l(t_f)$ is the Keplerian model to fit, similar to (2) but with correlated noise terms included. Maximization of this likelihood function allows signal detection to be performed and probabilities can be calculated directly from the log-likelihood values. Although in practice this method is slower than the aforementioned methods, it has the desired effect of allowing multiple signals to be detected at the same time (i.e., a global model approach), and it also means the model can include correlated noise components, along with the white noise component(s). Therefore, given the continuing increase in computer processing power, the extra information and flexibility of ML periodograms outweigh the inefficiency of its application to real radial velocity data. However, as with all model fitting methods, one must be careful not to overfit the data by adding unnecessary terms to the applied model, which is where proper model comparison statistical tests should be applied.

Bayesian analysis

Like the ML approach, Bayesian analysis applies a global model to the data, including correlated noise components, and assesses the parameter space using Markov chains (e.g., [15), where the model is assessed by covering a given frequency/period domain. The maximum of the posterior density distribution can be used to detect a signal in the data (e.g., [16] and [17])

$$m_f, d = \gamma_d + \gamma t_f + F_k(t_f) + \varepsilon_f, d + \sum_{z=1}^{q} c_{zd}\xi_z, f, d + \sum_{z=1}^{p} \phi_z, d \exp\left\{\frac{t_{f-z} - t_f}{\tau_d}\varepsilon_z, d\right\}. \quad (4)$$

Here model m for a given Keplerian k and velocity data point f, previous measurement z, and data set d can be described by the Keplerian model as a function of time ($F(t)$), a systemic offset velocity γ, a linear trend as a function of time γt, a Gaussian noise model to describe the random noise ε, a red noise component described by a moving average (MA) model with exponential smoothing (parameters ϕ and τ), and a set of linear correlations c with activity indicators that parameterise the activity state of the star at the time of the observation ξ. The Bayesian approach is the least efficient of these signal detection methods, since long chains are required to properly search the multidimensional parameter space in a robust manner. However, currently this method is the most flexible, allowing the user to assess the parameter space in many different ways. It also allows visualization of the full parameter space after the chains are complete, meaning nonlinear correlations between parameters can be scrutinized. Finally, this method was shown to be the most robust signal detection and false-positive suppression method currently used, given the results of an International Challenge (Extreme Precision Radial Velocities, Yale 2015) [30] issued to the radial velocity planet detection community.

MMSE-based method

In [18]–[20], the independent sinusoidal components in nonuniformly sampled radial velocity data are determined by means of the MMSE method or its direct extension, the ML estimation scheme. According to [19], significance tests are employed to filter out the parasitic solutions appearing on the way. In [18], the MMSE-based method applies a trellis-based optimal global search and returns the optimal number of sinusoidal components including their frequencies, phases, and amplitudes. This technique employs the MMSE criterion as an objective function in all the analysis.

If C_i is the ith sinusoidal component, and N_C is their number, each component may be written in the form $C_i = (\omega_i, a_i, \phi_i)$, where ω_i, a_i, ϕ_i are the frequency, amplitude, and phase of the ith component, respectively.

The MMSE technique tries to find the set $S = \{(\omega_i, a_i, \phi_i)\}_{i=1}^{N_C}$ that minimizes the mean square error between the original signal and S by optimizing ω_i, a_i, and ϕ_i of each component [18].

First, the target frequency bandwidth is divided into K_ω levels. Each level ω_k is represented by $\omega_k = \pi \times k/K_\omega$, where $1 \le k \le K_\omega$. For each ω_k an optimal amplitude and phase, $a_{\omega_k}, \phi_{\omega_k}$ are obtained by performing an MMSE-based Fourier analysis: for each ω_k, a_{ω_k} and ϕ_{ω_k} are optimized to minimize the mean square error between the original signal and the components $a_k \cos(\omega_k t + \phi_{\omega_k})$.

The number of components to analyze, N, is then estimated for all the frequencies having local minimum of MMSE values and/or higher amplitudes with respect to a defined threshold. Therefore, a subset $S_{\min} = \{(\omega_i, a_{\omega_i}, \phi_{\omega_i})\}$ is constructed out of the set S_P, which includes only these components. Next, a neighborhood band V_i is defined for each component, C_i in $S_{\min}$ as, $V_i = \{(\omega, a_\omega, \phi_\omega) \in S_P / \omega \in [\omega_i - \delta, \omega_i + \delta]\}$,

$$(\hat{\omega}_i^j, \hat{a}_i^j, \hat{\phi}_i^j)_{i=1}^{N_C} = \arg \min_{(\omega_i^j, a_i^j, \phi_i^j) 1 \le i \le N_C} \sum_{t=1}^{T} \left(x(t) - \sum_{i=1}^{N_C} a_i^j \cos(\omega_i^j t + \phi_i^j) \right)^2 \quad (5)$$

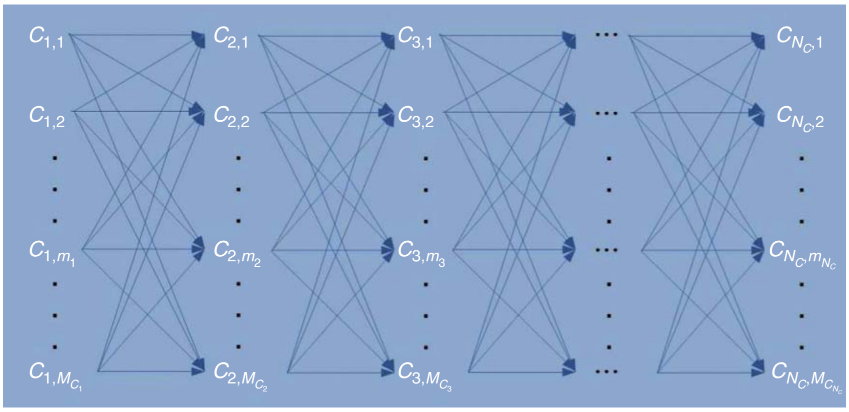


FIGURE 2. A trellis diagram representing all the possible combinations of components and their local neighborhoods [18].

where δ defines half of the neighborhood band around each component and set to some value that incorporates all the significant components around the selected peaks of S_P. Hence, for each component $C_i(\omega_i, a_{\omega i}, \phi_{\omega i})$ there are M_{C_i} candidates, where M_{C_i} is the cardinality of V_i.

Subsequently, a trellis analysis for all of the possible combinations of components and their local neighborhoods is performed, as schematically shown in Figure 2. Now, all possible combinations of candidates *NC* is evaluated, where $N_C = \{1,..,N\}$. For each value of N_C, let $A^j = \{C_1^j, C_2^j, ..., C_{N_C}^j\}$ be a set of triplets for one of the possible combinations, where j is from 1 to $N!/N_C!(N-N_C)!$ and C_i^j corresponds to the ith component in the jth combination. The corresponding set of neighborhoods is $V_{A^j} = \{V_1^j, V_2^j, ..., V_{N_C}^j\}$, where $V_i^j = (\omega_i^j, a_i^j, \phi_i^j)$ denotes the neighborhood of candidate components C_i^j. The optimal set of A^j, $\hat{A}^j$, for a specific N_C value, is the one associated with the lowest MMSE and corresponds to (5), shown in the box at the bottom of the previous page [18], where $(\hat{\omega}_i^j, \hat{a}_i^j, \hat{\phi}_i^j) \in V_i^j$, $1 \leq \text{i} \leq N_C$, providing an optimal set of triplets for each N_C. Finally, the optimal set of triplets having the global minimum MMSE is selected at which its length, defines the number of the most important sinusoidal components in the nonuniformly sampled signal, while its elements are their frequencies, amplitudes, and phases, respectively. It is worth mentioning that the problem of order selection has also been addressed by using statistical significance analysis [20] and extreme value theory [21].

Figure 3 illustrates, as an example, the Keck and the High Accuracy Radial Velocity Planetary Searcher radial velocity data of the M-dwarf planet host star GJ876 and the corresponding LS and MMSE periodograms. GJ876 is known to host a system of planets that contains at least two short-period gas giants [18]. Signals can be searched for using the gradient-based approach that starts by searching for one signal only, and when one is detected it is then subtracted out of the data and a new search is made using the residuals all over again by treating them as an independent time series from the original observed data. This process is then repeated until the noise floor of the data is reached. By applying this method, the following signals [with periods in days (d)] were detected with the MMSE method [18]: 61.03 d, 30.23 d, 15.04 d, 1.94 d, 10.01 d, and 124.69 d. This system was chosen because the two large-amplitude signals could be detected in both halves of the time series separately. The MMSE and trellis technique allows studying the phase of the detected signals as a function of time, showing that the phase difference between both planets is stable over the length of the time series and therefore adding weight to the reality of these signals. This analysis shows the power of this method over previous periodogram techniques, such as the LS method, that gives no information on the signal parameters other than the frequency. However, phase variations with time for the 1.94 d, 10.01 d, and 15.04 d signals were found, which could cast doubt on the origin of these signals as being from orbiting planets. This was consistent with previous Newtonian integrational methods. This highlights that

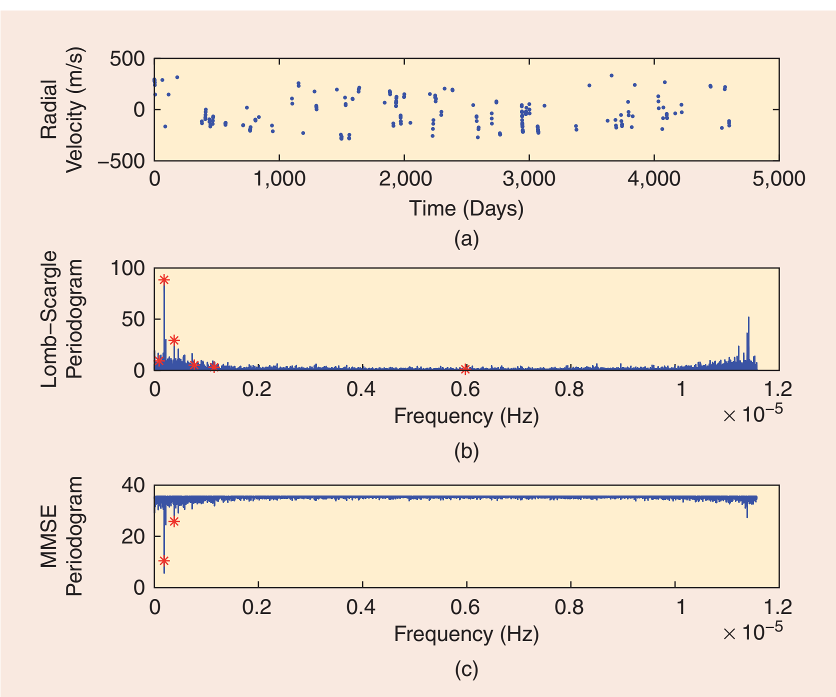


FIGURE 3. Radial velocity data of (a) star GJ876, (b) the LS and MMSE periodograms, and (c) the planets initially detected shown by asterisks.

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

the MMSE method provides the flexibility of further validating the authenticity of the signals, runs a global search for all the signals in the entire data, and outputs the frequency, phase, and amplitude of the signals. Nevertheless, the MMSE and trellis search method does not include any correlated red noise model, whereas the ML and Bayesian analysis do, and since correlated noise does indeed appear to be a very important part of high-precision radial velocity analysis at the ~m/s level, as mentioned previously, avenues to test here would be the application of Gaussian processes (e.g., [22] and [23]), MAs [24], among others applied in the field and those yet to be tested.

Statistical significance of signal detection

For any signal detection method, a robust statistical validation should be made of any detected signal, likely calculating the probability directly that the signal could be due to random noise fluctuations. For instance, for the LS method [6], the probability of a signal at any given frequency follows an exponential distribution, where the larger the number of frequencies sampled, the larger the probability that a matching frequency is found. They defined a false alarm probability (FAP) analytically, such that one can determine the probability of any given frequency being real, solely based on the signal's measured power and the number of frequencies sampled.

Given the nature of problems in astronomy, the deviations from normality, the excess noise in measurements, etc., it has become normal to instead calculate FAPs directly from the data using nonparametric statistical methods, like bootstrap analysis, for example (see [25] and [26]). Bootstrapping is performed by scrambling the radial velocity data with replacement, maintaining the time stamps, then reconstructing the periodogram and selecting the highest peak. Each of the strongest peaks are recorded from a series of 10,000 or more independent trials, and the total number of peaks found to be stronger than the observed peak power provides a direct measure of the FAP, or how much such a power can arise from random chance.

Finally, the ML periodograms and Bayesian method allow probabilities to be drawn directly from the data. We previously discussed that the ML method allows probabilities to be calculated for each frequency as part of the methodology. For the Bayesian approach, statistical comparison tests can be performed to assess if certain models are better suited to the data in comparison to flat noise models, for instance. It is common to calculate the Bayes factors to evaluate if one model is statistically favored over another, since this method is based on marginalization of the likelihood, a process that naturally applies a penalty to models with increasing complexity (so-called Occam's penalty; see [27]). In fact, teams who employ these types of methods are known to favor certain models over others, only if they are at least 10,000 times more probable (e.g., [24]).

Potential directions for future research

We want to detect exo-Earths so future directions for the radial velocity method are better calibrations. One big avenue of research is the implementation of laser comb technology, which recent tests have told us will allow velocity stability at the centimeter/second level, necessary for the discovery of Earth-like worlds. Furthermore, some areas of stellar astrophysics needs to be better understood, particularly the impact of stellar activity on radial velocity measurements. All of the methods reviewed in this article have the potential to be optimized to further enhance the detection results. We need to better model the impact of magnetic activity on radial velocities. In fact, this impacts transits, transit timing variations, and astrometry measurements. New signal processing methods for signal enhancement and red noise modeling and removal also need to be investigated (e.g., see [28] and [29]).

Acknowledgments

Our work was partially funded by the CONICYT-PIA project ACT 1120 "Center for Multidisciplinary Research on Signal Processing," Chile. James Stewart Jenkins also acknowledges funding from Fondecyt grant 1161218 and BASAL CATA PFB-06. Muhammad Salman Khan's work was funded by the CONICYT-PIA project ACT 1120.

Authors

Muhammad Salman Khan (salmankhan@uetpeshawar.edu.pk) is an assistant professor in the Department of Electrical Engineering at the University of Engineering and Technology, Peshawar, Jalozai campus, Pakistan. He was a postdoctoral researcher in the Department of Electrical Engineering, Universidad de Chile, Santiago between 2013 and 2015. His research interests include signal processing, pattern recognition, machine learning, and big data. He is a Member of the IEEE and the IEEE Signal Processing Society.

James Stewart Jenkins (jjenkins@das.uchile.cl) is an assistant professor in the Departamento de Astronomia, Universidad de Chile, Santiago. His research interests are mainly focused on the search for extrasolar planets using the radial velocity method. In particular, he employs correlated noise models to detect low-mass planets orbiting nearby stars and was a member of the team who discovered the habitable zone terrestrial planet Proxima Centauri b.

Nestor Becerra Yoma (nbecerra@ing.uchile.cl) is a full professor in the Department of Electrical Engineering, Universidad de Chile, Santiago. Since August 2016, he has been a visiting professor at Carnegie Mellon University, Pittsburgh, Pennsylvania. At the Universidad de Chile, he started the Speech Processing and Transmission Laboratory to carry out research on speech technology applications on human-robot interaction, Internet, and telephone. He was also the director and principal investigator of the Center for Multidisciplinary Research on Signal Processing project. He is a member of the IEEE Signal Processing Society and the International Speech Communication Association.

References

[1] M. Mayor, M. Marmier, C. Lovis, S. Udry, D. Segransan, F. Pepe, W. Benz, J.-L. Bertaux, F. Bouchy, X. Dumusque, G. Lo Curto, C. Mordasini, D. Queloz,

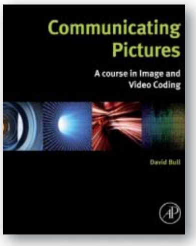
(continued on page 115)

BOOK DIGEST

EDITORS' INTRODUCTION

Books focusing on signal processing are constantly published by academic publishers and researchers. To enhance the visibility of new signal processing books and to inform our readers of recently published books in a timely fashion, *IEEE Signal Processing Magazine* (*SPM*) launched the first "Book Digest" column in its January 2016 issue. Different from the "Book Review" column, which requires capable reviewers and takes a lengthy time to complete a review, the "Book Digest" column provides a list of books with a concise summary for each one. Books are selected by a pool of senior editors and are based on criteria such as timeliness of the topic, track record of the authors, training materials for students, and signal processing focus. If an expert volunteer is available to review a book that has a high impact on signal processing, a review will be considered for publication in the "Book Review" column. Should you have any comments or wish to have your book considered for publication in this column, do not hesitate to contact Kenneth Lam (enkmlam@polyu.edu.hk), *SPM*'s area editor, columns and forums, or Danilo Mandic (d.mandic@imperial.ac.uk), *SPM*'s associate editor, "Book Digest" and "Book Review" columns.

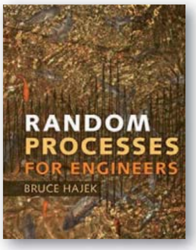
David R. Bull. *Communicating Pictures: A Course in Image and Video Coding.* Academic Press/Elsevier, Year: 2014, ISBN: 9780124059061.



Communicating Pictures starts with a unique historical perspective of the role of images in communications and then builds on this to explain the applications and requirements of a modern video coding system. It draws on the author's extensive academic and professional experience of signal processing and video coding to deliver a text that is algorithmically rigorous, yet accessible, relevant to modern standards, and practical. It offers a thorough grounding in visual perception and demonstrates how modern image and video compression methods can be designed to meet the rate-quality performance levels demanded by today's applications, networks, and users.

With this book you will learn: 1) practical issues when implementing a codec, such as picture boundary extension and complexity reduction, with particular emphasis on efficient algorithms for transforms, motion estimators and error resilience; 2) conflicts between conventional video compression, based on variable length coding and spatiotemporal prediction, and the requirements for error resilient transmission; 3) how to assess the quality of coded images and video content, both through subjective trials and by using perceptually optimised objective metrics; and 4) features, operation, and performance of the state-of-the-art high-efficiency video coding standard.

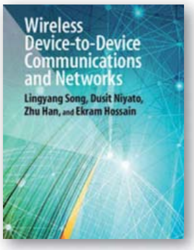
Bruce Hajek. *Random Processes for Engineers.* Cambridge University Press, Year: 2015, ISBN: 9781107100121.



This engaging introduction to random processes provides students with the critical tools needed to design and evaluate engineering systems that must operate reliably in uncertain environments.

A brief review of probability theory and a real analysis of deterministic functions set the stage for understanding random processes, while the underlying measure theoretic notions are explained in an intuitive, straightforward style. Students will learn to manage the complexity of randomness through the use of simple classes of random processes, statistical means and correlations, asymptotic analysis, sampling, and effective algorithms. Key topics covered include calculus of random processes in linear systems, Kalman and Wiener filtering, hidden Markov models for statistical inference, the estimation maximization algorithm, and an introduction to martingales and concentration inequalities. Understanding of the key concepts is reinforced through more than 100 worked examples and 300 thoroughly tested homework problems.

Lingyang Song, Dusit Niyato, Zhu Han, and Ekram Hossain. *Wireless Device-to-Device Communications and Networks.* Cambridge University Press, Year: 2015, ISBN: 9781107063570.



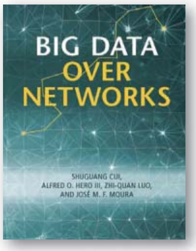
Covering the fundamental theory together with the state of the art in research and development, this practical guide provides the techniques needed to design, analyze, and optimize device-to-device (D2D) communications in wireless networking.

With an ever-increasing demand for higher-data-rate wireless access, D2D communication is set to become a key

Digital Object Identifier 10.1109/MSP.2016.2618578
Date of publication: 11 January 2017

feature supported by next-generation cellular networks. This book introduces D2D-based wireless communications from the physical-, media access control-, network-, and application-layer perspectives, providing all the key background information before moving on to discuss real-world applications as well as potential future developments. Key topics are discussed in detail, such as dynamic resource sharing (e.g., of spectrum and power) between cellular and ad hoc D2D communications to accommodate larger volumes of traffic and provide better service to users. Readers will understand the practical challenges of resource management, optimization, security, standardization, and network topology, and learn how the design principles are applied in practice.

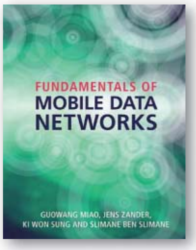
Shuguang Cui, Alfred O. Hero III, Zhi-quan Luo, and José M.F. Moura (Editors). Big *Data over Networks*. Cambridge University Press, Year 2016, ISBN: 9781107099005.



Utilizing both key mathematical tools and state-of-the-art research results, this text explores the principles underpinning large-scale information processing over networks and examines the crucial interaction between big data and its associated communication, social, and biological networks.

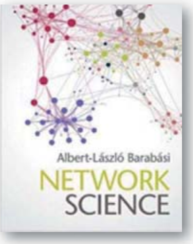
Written by experts in the diverse fields of machine learning, optimization, statistics, signal processing, networking, communications, sociology, and biology, this book employs two complementary approaches: 1) analyzing how the underlying network constrains the upper layer of collaborative big data processing and 2) examining how big data processing may boost performance in various networks. Unifying the broad scope of the book is the rigorous mathematical treatment of the subjects, which is enriched by in-depth discussion of future directions and numerous open-ended problems that conclude each chapter. Readers will be able to master the fundamental principles for dealing with big data over large systems, making it essential reading for graduate students, scientific researchers, and industry practitioners alike.

Guowang Miao, Jens Zander, Ki Won Sung, and Slimane Ben Slimane. *Fundamentals of Mobile Data Networks*. Cambridge University Press, Year: 2016, ISBN: 9781107143210.



This unique text provides a comprehensive and systematic introduction to the theory and practice of mobile data networks. Covering basic design principles as well as analytical tools for network performance evaluation, and with a focus on system-level resource management, you will learn how state-of-the-art network design can enable you to flexibly and efficiently manage and trade off various resources such as spectrum, energy, and infrastructure investments. Topics covered range from traditional elements such as medium access, cell deployment, capacity, handover, and interference management, to more recent cutting-edge topics such as heterogeneous networks, energy and cost-efficient network design, and a detailed introduction to long-term evolution (4G). Numerous worked examples and exercises illustrate the key theoretical concepts and help you put your knowledge into practice, making this an essential resource whether you are a student, researcher, or practicing engineer.

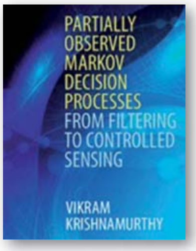
Albert-László Barbási. *Network Science*. Cambridge University Press, Year: 2016, ISBN: 9781107076266.



Networks are everywhere, from the Internet, to social networks, and the genetic networks that determine our biological existence. Illustrated throughout in full color, this pioneering textbook, spanning a wide range of topics from physics to computer science, engineering, economics and social sciences, introduces network science to an interdisciplinary audience.

From the origins of the six degrees of separation to explaining why networks are robust to failures and fragile to attacks, the author explores how viruses like Ebola and H1N1 spread and why it is that our friends have more friends than we do. Using numerous real-world examples, this innovative text includes clear delineation between undergraduate- and graduate-level material. The mathematical formulas and derivations are included within advanced topics sections, enabling use at a range of levels. Extensive online resources, including films and software for network analysis, make this a multifaceted companion for anyone with an interest in network science.

Vikram Krishnamurthy. *Partially Observed Markov Decision Processes: From Filtering to Controlled Sensing*. Cambridge University Press, Year: 2016, ISBN: 9781316471104.



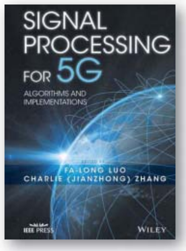
Covering formulation, algorithms, and structural results and linking theory to real-world applications in controlled sensing (including social learning, adaptive radars, and sequential detection), this book focuses on the conceptual foundations of partially observable Markov decision processes (POMPDs). It emphasizes structural results in stochastic dynamic programming, enabling graduate students and researchers in engineering, operations research, and economics to understand the underlying unifying themes without getting weighed down by mathematical technicalities. Bringing together research from across the literature, the book provides an introduction to nonlinear filtering followed by a systematic development of stochastic dynamic programming, lattice programming, and reinforcement learning for POMDPs.

The abstraction of POMDPs becomes alive with applications. This book contains several examples starting from target tracking in Bayesian filtering to optimal search, risk measures, active sensing, adaptive radars, and social learning.

The supplement of this book contains errata and problem sets and can be accessed via: http://www.cambridge.org/gb/academic/subjects/engineering/communications-and-signal-processing/partially-observed-markov-decision-processes-filtering-controlled-sensing?format=HB.

Fa-Long Luo and Charlie Jianzhong Zhang (Editors). *Signal Processing for 5G: Algorithms and Implementations.* Wiley-IEEE Press, Year: 2016, ISBN: 9781119116462.

Signal processing techniques have played the most important role in wireless



communications since the second generation of cellular systems. It is anticipated that new techniques employed in fifth-generation (5G) wireless networks will not only improve peak service rates significantly but also enhance capacity, coverage, reliability, low-latency, efficiency, flexibility, compatibility, and convergence to meet the increasing demands imposed by applications such as big data, cloud service, machine-to-machine, and mission-critical communications.

This book is a comprehensive and detailed guide to all signal processing techniques employed in 5G wireless networks. Uniquely organized into four categories, "New Modulation and Coding, "New Spatial Processing," "New Spectrum Opportunities," and "New System-Level Enabling Technologies," it covers everything from network architecture, physical layer (down-link and up-link), protocols and air interface, to cell acquisition, scheduling and rate adaption, access procedures, and relaying to spectrum allocations. All technology aspects and major roadmaps of global 5G standard development and deployments are included in the book. SP

LECTURE NOTES (continued from page 103)

Conclusions

This article introduces a new paradigm of PP techniques—CP—which represents a dimension-reduced subspace approach to PP machine learning. Built upon the information and estimation theory, CP methods tackle joint optimization over feature/utility/privacy spaces. This leads to several eigen-system-based subspace methods, including PCA, DCA, and DUCA. To confirm the theoretical analysis, we have conducted experimental studies on various DIP and FR problems. The latter also demonstrates possible real-world applications of the proposed CP methodology.

Acknowledgments

This material is based upon work supported in part by the Brandeis Program of DARPA and the Space and U.S. Naval Warfare System Center Pacific (SSC Pacific) under contract 66001-15-C-4068. I would like to thank Prof. Morris J. Chang (ISU) and Prof. Peiyuan Wu, Yuan Zhou, Ying Li, Dr. Shibiao Wan, Thee Chanyaswad, Mert Al, Chang Chang Liu, and Artur Filipowicz (Princeton University) for invaluable discussions and assistance.

Author

S.Y. Kung (kung@princeton.edu) is a professor in the Department of Electrical Engineering at Princeton University, New Jersey. His research areas include machine learning, compressive privacy, data mining and analysis, statistical estimation, system identification, wireless communication, very-large-scale integration (VLSI) array processors, genomic signal processing, and multimedia information processing. He was a founding member of several technical committees of the IEEE Signal Processing Society. He served as a member of the Board of Governors of the IEEE Signal Processing Society (1989–1991). He has been the editor-in-chief of *Journal of VLSI Signal Processing Systems* since 1990. He has received multiple awards and recognitions. He is a Life Fellow of the IEEE.

References

[1] C. Dwork, K. Kenthapadi,F, McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *Advancves in Cryptology-EUROCRYPT*. New York: Springer, 2006.

[2] F. du Pin Calmon and N. Fawaz, "Privacy against statistical inference," in *Proc. Allerton Conf. Communication, Control, and Computing*, 2012, pp. 1401–1408.

[3] S. Y. Kung, *Kernel Methods and Machine Learning*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[4] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," in *Proc. 37th Allerton Conf. Communication and Computation*, 1999.

[5] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. 2nd ed. New York: Wiley, 2011.

[6] S. Y. Kung, "Discriminant component analysis for privacy protection and visualization of big data," *Multimed. Tools Appl.*, 2015, doi:10.1007/s11042-015-2959-9.

[7] S. Y. Kung, T. Chanyaswad, J. Morris Chang, and P. Wu, "Collaborative PCA/DCA learning methods for compressive privacy," *ACM Trans. Embed. Comput. Syst.* (Special Issue on Effective Divide-and-Conquer, Incremental, or Distributed Mechanisms of Embedded Designs for Extremely Big Data in Large-Scale Devices), to be published.

SP

DATES AHEAD

Please send calendar submissions to:
Dates Ahead, Att: Jessica Barragué, E-mail: j.barrague@ieee.org

2017

MARCH

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)
5–9 March, New Orleans, Louisiana, USA.
General Chair: Magdy Bayoumi
URL: http://www.ieee-icassp2017.org/

APRIL

IEEE International Symposium on Biomedical Imaging (ISBI)
18–21 April, Melbourne, Australia.
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. Sego
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/

Digital Object Identifier 10.1109/MSP.2016.2636082
Date of publication: 11 January 2017

IMAGE LICENSED BY INGRAM PUBLISHING

ICASSP 2017 will be held in New Orleans, Louisiana, 5-9 March.

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/

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 Distante and Larry S. Davis
URL: www.avss2017.org

SEPTEMBER

IEEE International Conference on Image Processing (ICIP)
17–20 September, Beijing, China.
General Chairs: Xinggang Lin, Anthony Vetro, and Min Wu
URL: http://2017.ieeeicip.org/

OCTOBER

IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)
15–18 October, New Paltz, New York.
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.eee.strath.ac.uk/

NOVEMBER

5th 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

IN THE SPOTLIGHT *(continued from page 116)*

ADOBE STOCK

much like data science, many of the disciplines within signal processing are fundamentally about looking for correlations and dependencies in data to effectively make decisions.

That is not to say that signal processing is the same as data science. Perhaps the most notable difference between the past era and the new era of data science and big data is the tearing down of boundaries associated with how data is produced and accessed. Big data is fundamentally heterogeneous, involving data from a vast collection of sources that report data of various modalities for analysis. Whereas the previous generation of scientific discovery involved scientists conducting (and planning) experiments to intentionally measure specific data for the purpose of discovery, oftentimes big data involves the opportunistic sharing of data from nonvetted sources often provided in unstructured representations. Thus, the new era of big data and data analytics will likely lead to new engineered systems that utilize data from sources previously unknown to the engineer and application developer. In short, the new era will have more data than you could ever dream.

But, as we move forward in this new era, we need to relish in the opportunities it will provide, yet also retain an appropriate level of caution. The promise of being able to analyze large amounts of data to find a cure for cancer, integrate infrastructure and vehicle sensor data to allow for automated driving and more efficient transportation, or the potential to analyze the data being generated by the broad collection of astronomical observatories to discover new stellar phenomena are certainly fantastic and truly important to society. We would not be able to make such advancements or build new systems without the emergence of this new field of data science. However, we must be careful as this explosion of data and data science could take on a life of its own. Regardless of whether you are a scientist, mathematician, engineer, or in some other profession, you were likely raised with the "scientific method" drummed into you like a mantra. We've all grown up in an era of slow, methodical research and development. In fact, one way of looking at both the scientific method and the engineering design process is that it leads to implicit practice of quality control—almost bordering on pessimism and overt caution.

> **As this new era of big data and data science unfolds, let us issue a challenge to scientists, engineers, and signal processors to establish new forms of collaboration.**

Big data will often involve others unintentionally conducting experiments for the data scientist. The allure of hunting through more and more data to find patterns without vetting that data is dangerous. Data science will have some growing pains, especially as the vast amount of data being examined guarantees that data will be haphazardly analyzed and spurious correlations will be proclaimed as scientific truths. Data science will need quality control.

And this is where the signal processing community can advance big data and data science. Over the years, the signal processing community has carefully built up a sophisticated toolbox full of algorithms designed to analyze data, as well as the deep understanding of when and how to use these algorithms, and how they can be made to work efficiently. Signal processors are a mixed breed of statisticians crossed with control theorists crossed with computer engineers who have, over the decades, folded performance assurance into their algorithms to ensure that video looks good after compressed, targets are accurately tracked, and tumors can be effectively classified with low rates of false alarm and missed detections.

Hence, as this new era of big data and data science unfolds, let us issue a challenge to scientists, engineers, and signal processors to establish new forms of collaboration: To the data scientists, reach out and ask a signal processor whether they know of any signal processing tools that might work on your data. To the signal processor, find the scientists and engineers who are making the next wave of data and offer your services. Now more than ever is the time for those engaged in signal processing to reach across the boundaries of technical fields and contribute their tools to the analysis of the vast amounts of data that are being generated everywhere. Signal processing has had a fantastic record of success, and, as we move to this new world of data treasure hunting, signal processing can ensure the success of data science—ensuring that the hidden correlations one finds are truly golden treasures and not spurious pyrite counterfeits.

Author

Wade Trappe (trappe@winlab.rutgers.edu) is a professor in the Electrical and Computer Engineering Department at Rutgers University, New Jersey, and associate director of the Wireless Information Network Laboratory (WINLAB), where he directs WINLAB's research in wireless security. He coauthored the textbook *Introduction to Cryptography with Coding Theory* as well as several monographs on wireless security, including *Securing Wireless Communications at the Physical Layer* and *Securing Emerging Wireless Systems: Lower-Layer Approaches*. He was an editor of *IEEE Transactions on Information Forensics and Security*, *IEEE Signal Processing Magazine*, and *IEEE Transactions on Mobile Computing*. He is an IEEE Fellow.

SP

APPLICATIONS CORNER

(continued from page 109)

and N. C. Santos "The HARPS search for southern extra-solar planets xxxiv. occurrence, mass distribution and orbital properties of super-Earths and Neptune-mass planets," *Astron. Astrophys*., Sept. 2011 no. arXiv:1109.2497.

[2] G. Anglada-Escud, et al. "Two planets around Kapteyn's star: a cold and a temperate super-Earth orbiting the nearest halo red dwarf," *Mon. Not. R. Astron. Soc*., vol. 443, no. 89, pp. L89-L93, 2014.

[3] J. S. Jenkins, et al. "New planetary systems from the Calan-Hertfordshire extrasolar planet search and core accretion mass limit," *Mon. Not. R. Astron. Soc*., Oct. 2015.

[4] A. Baranne, et al. "ELODIE: A spectrograph for accurate radial velocity measurements," *Astron. Astrophys*., vol. 119, pp. 373–390, Oct. 1996.

[5] J. S. Jenkins, H. R. A. Jones, Y. Pavlenko, D. J. Pinfield, J. R. Barnes, and Y. Lyubchik "Metallicities and activities of southern stars," *Astron. Astrophys*., vol. 485, no. 2, pp. 571–584, 2008.

[6] J. D. Scargle, "Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data," *Astrophy. J*., vol. 263, pp. 835–853, Dec. 1982.

[7] G. L. Bretthorst, *Bayesian Spectrum Analysis and Parameter Estimation* (Lecture Notes in Statistics, vol. 48). New York: Springer-Verlag, 1988.

[8] S. Ferraz-Mello, "Estimation of periods from unequally spaced observations," *Astron. J*., vol. 86, no. 4, pp. 619–624, 1981.

[9] A. Cumming, "Detectability of extrasolar planets in radial velocity surveys," *Mon. Not. R. Astron. Soc*., vol. 354, no. 4, pp. 1165–1176, 2004.

[10] M. Zechmeister and M. Kürster, "The generalised Lomb–Scargle periodogram," *Astron. Astrophys*., vol. 496, no. 2, pp. 577–584, 2009.

[11] R. V. Baluev, "Keplerian periodogram for doppler exoplanet detection: optimized computation and analytic significance thresholds," *Mon. Not. R. Astron. Soc*., vol. 446, no. 2, pp. 1478–1492, 2015.

[12] P. Lenz and M. Breger, "Period04: A software package to extract multiple frequencies from real data," in *The A-Star Puzzle: IAUS 224*. Cambridge, U.K.: Cambridge Univ. Press, 2004, pp. 786–790.

[13] R. V. Baluev, "A radial-velocity time-series analysis tool facilitating exoplanets detection, characterization, and dynamical simulations," *Astron. Comput*., vol. 2, pp. 18–26, Aug. 2013.

[14] G. Anglada-Escud, M. Tuomi, et al. "A dynamically-packed planetary system around GJ 667C with three super-Earths in its habitable zone," *Astron. Astrophys*., vol. 556, pp. 126–150, June 2013.

[15] H. Haario, E. Saksman, and J. Tamminen, "An adaptive metropolis algorithm," *Bernoulli*, vol. 7, no. 2, 2001.

[16] M. Tuomi, "A new cold sub-Saturnian candidate planet orbiting GJ 221," *Mon. Not. R. Astron. Soc*., vol. 440, pp. L1–L5, Jan. 2014.

[17] J. Jenkins and M. Tuomi, "The curious case of HD 41248. A pair of static signals buried behind red noise," *Astrophys. J*., vol. 794, no. 2, 2014.

[18] J. S. Jenkins, N. B. Yoma, P. Rojo, R. Mahu, and J. Wuth, "Improved signal detection algorithms for unevenly sampled data. Six signals in the radial velocity data for GJ876," *Mon. Not. R. Astron. Soc*., vol. 441, pp. 2253–2265, Mar. 2014.

[19] R. V. Baluev, "Detecting multiple periodicities in observational data with the multifrequency periodogram-II. Frequency decomposer, a parallelized time-series analysis algorithm," *Astron.Comput*., vol. 3–4, pp. 50–57, Nov.–Dec. 2013.

[20] R. V. Baluev, "Detecting multiple periodicities in observational data with the multifrequency periodogram–I. Analytic assessment of the statistical significance," *Mon. Not. R. Astron. Soc*., vol. 436, pp. 807–818, Aug. 2013.

[21] F. Talebi and T. Pratt, "Model order selection for complex sinusoids in the presence of unknown correlated Gaussian noise," *IEEE Trans, Signal Process*., vol. 63, no. 7, pp. 1664–1674, 2015.

[22] R. V. Baluev, "The impact of red noise in radial velocity planet searches: Only three planets orbiting GJ 581?" *Mon. Not. R. Astron. Soc*., vol. 429, no. 3, pp. 2052–2068, 2012.

[23] V. Rajpaul1, S. Aigrain, M. A. Osborne, S. Reece, and S. Roberts, "A Gaussian process framework for modelling stellar activity signals in radial velocity data," *Mon. Not. R. Astron. Soc*., vol. 452, no. 3, pp. 2269–2291, 2015.

[24] M. Tuomi, et al. "Signals embedded in the radial velocity noise. Periodic variations in the τ Ceti velocities," *Astron. Astrophys*., vol. 551, pp. A79, Dec. 2012.

[25] B. Efron, "Bootstrap methods: Another look at the jacknife," *Ann. Stat*., vol. 7, no 1, pp. 1–126, 1979.

[26] J. S. Jenkins, et al. "A hot Uranus orbiting the super metal-rich star HD77338 and the metallicity-mass connection," *Astrophys. J*., vol. 766, no. 2, 2013.

[27] P. C. Gregory, "Bayesian re-analysis of the Gliese 581 exoplanet system," *Mon. Not. R. Astron. Soc*., vol. 415, pp. 2523–2545, Jan. 2011.

[28] J. A. Carter and J. N. Winn, "Parameter estimation from time-series data with correlated errors: A wavelet-based method and its application to transit light curves," *Astrophys. J*., vol. 704, pp. 51–67, Oct. 2009.

[29] B. Placek, K. H. Knuth, D. Angerhausen, and J. M. Jenkins, "Characterization of Kepler-91b and the investigation of a potential trojan companion using EXONEST," *Astrophys. J*., vol. 814, pp. 147, Dec. 2015.

[30] X. Dumusque, "Radial velocity fitting challenge-I: Simulating the data set including realistic stellar radial-velocity signals," *Astron. Astrophys*., vol. 593, vol. 5, 2016.

SP

Wade Trappe

Data Treasure Hunters: Science Expanding to New Frontiers

EDITOR'S NOTE

In June 2016, the IEEE Signal Processing Society (SPS) launched an SPS blog website, which provides a nontechnical supplement to highly technical signal processing topics. There are seven SPS blogs so far, and these blogs help students and the general public to become more aware of signal processing. We selected the first SPS blog in the web series, written by Wade Trappe, for this issue's "In the Spotlight" column. We hope you enjoy reading it. You can find more SPS blogs by visiting http://signalprocessingsociety.org/publications-resources/blog. Should you have any suggestions regarding the SPS blogs, please do not hesitate to contact *IEEE Signal Processing Magazine*'s Area Editor, Columns and Forums Kenneth Lam (enkmlam@polyu.edu.hk) and SPS Membership and Content Administrator Jessica Perry (jessica.perry@ieee.org).

Science and engineering are rapidly heading toward a major culture change—a change in how we think about data.

This change is already happening, and it will be dramatic and exciting! It will completely change how most of us think about data and how we tackle science and engineering problems. With it will come a flood of new discoveries—advances in the sciences and in new technologies—that were never before possible. What is this revolution? How did we get here? Where is it going, and how is signal processing involved?

The short answer is that we are entering an era of treasure hunting. Rather than digging through dirt like archaeologists looking for ancient artifacts, the future will involve digging through data.

Digital Object Identifier 10.1109/MSP.2016.2619918
Date of publication: 11 January 2017

We are experiencing an explosion in the amount of data available for scientists and engineers to do their jobs. The world around us is becoming increasingly "connected" as communication technologies have proliferated and the costs of digital data storage have plummeted. Seemingly mundane items and devices that never before had a bit or byte associated with them are now streaming a constant flow of data to data warehouses located in the cloud. Advancements in medical devices are leading to the emergence of miniaturized, nonintrusive medical sensors that will be integrated with communication technologies to report real-time glucose levels, monitor respiratory conditions, track immune responses, and allow for the analysis of a wide array of other data associated with the human condition. Meanwhile, scientific equipment is being aimed both out into space as well as deep into the Earth and across its ecosystems. Matching this explosion in data is a commensurate advance in computing: computing resources are now sophisticated enough to be able to perform immense amounts of computation on this data.

This is the emergence of the new field of big data and data analytics. Data scientists are the postmodern treasure hunters. They will reach into their toolboxes of algorithms and dig into data looking for hidden correlations, trying to find never-before-seen patterns with the hope of advancing the frontier of knowledge and supporting the development of new products.

The frank truth, though, is that data science isn't really new. Many technical fields have been performing analysis on large amounts of data before the term *big data* was ever coined. The signal processing community has been analyzing data since its inception. After all, what is the Fourier transform but a tool to find periodic phenomena in data? Or take a quick survey of papers over the past 25 years (or more), and you will find signal processing is involved in everything from analyzing geological data for oil discovery, to face recognition for domestic security, to processing genomic data and looking for patterns that indicate the onset of cancer. Signal processing was fundamental to advancements in multimedia processing and storage, and

(continued on page 114)

 1053-5888/17©2017IEEE

IEEE GlobalSIP

5th IEEE Global Conference on Signal and Information Processing

November 14 - 16, 2017, Montreal, Canada

IEEE Advancing Technology for Humanity

IEEE Signal Processing Society

http://www.ieeeglobalsip.org
2017.ieeeglobalsip.org

Call for Symposium Proposals

We invite Symposium proposals for the fifth IEEE Global Conference on Signal and Information Processing (GlobalSIP) which will be held in Montreal, Quebec, Canada on November 14-16, 2017. GlobalSIP is a flagship IEEE Signal Processing Society conference. It 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. Topics include but are not limited to:

- Signal and information processing for
  - communications and networks, including green communications
  - optical communications
  - forensics and security
  - finance
  - energy and power systems (e.g., smart grid)
  - genomics and bioengineering (physiological, pharmacological and behavioral)
  - neural networks, including deep learning
- Image and video processing
- Selected topics in speech processing and human language technologies
- Human machine interfaces
- Multimedia transmission, indexing, retrieval, and quality of experience
- Selected topics in statistical signal processing
- Cognitive communications and radar
- Graph-theoretic signal processing
- Machine learning
- Compressed sensing and sparsity aware processing
- Seismic signal processing
- Big data and social media challenges
- Hardware and real-time implementations
- Other (industrial) emerging applications of signal and information processing.

Symposium proposals should contain the following information: title; duration (e.g., full day or half day); paper length, acceptance rate; name, address, and a short CV (up to 250 words) of the organizers, including the technical chairs (if any); a 1-page or 2-page description of the topics to be addressed, including timeliness and relevance to the signal processing community; names of (potential) members of the technical program committee; invited speakers' name; a draft call for papers. Please pack everything together in a single pdf document. More detailed information can be found in GlobalSIP2017 Symposium Proposal Preparation Guide.

Proposed Timeline

- **Jan. 20, 2017**: Symposium proposals due
- **Jan. 25, 2017**: Symposium selection decision made
- **Feb. 1, 2017**: Call for Papers for accepted Symposia
- **May 15, 2017**: Paper submission due
- **June 30, 2017**: Notification of Acceptance
- **July 22, 2017**: Camera-ready paper due.

Digital Object Identifier 10.1109/MSP.2016.2636084

IEEE SIGNAL PROCESSING SOCIETY

Content Gazette

JANUARY 2017 ISSN 2167-5023

T-SP January 1 Vol. 65 #1
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7755888

T-SP January 15 Vol. 65 #2
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7762278

T-ASLP December Vol. 24 #12
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7755865

T-CI December Vol. 2 #4
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7736166

T-IP December Vol. 25 #12
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7744749

T-IFS December Vol. 11 #12
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7747716

T-MM December Vol.18 #12
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7744756

T-JSTSP December Vol. 10 #8
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7746026

T-SIPN December Vol. 2 #4
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7736190

T-SPL December Vol. 23 #12
http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7769278

http://www.ieee.org/publications_standards/publications/authors/author_ethics.html

IEEE Signal Processing Society

IEEE

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

IEEE Signal Processing Society

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

JANUARY 1, 2017 VOLUME 65 NUMBER 1 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

Second-Generation Curvelets on the Sphere http://dx.doi.org/10.1109/TSP.2016.2600506 .... *J. Y. H. Chan, B. Leistedt, T. D. Kitching, and J. D. McEwen* 5
Theoretical Bounds in Minimax Decentralized Hypothesis Testing http://dx.doi.org/10.1109/TSP.2016.2613072 .... *G. Gül and A. M. Zoubir* 15
A Signal-Space Aligned Network Coding Approach to Distributed MIMO http://dx.doi.org/10.1109/TSP.2016.2616335 .... *T. Yang, X. Yuan, and Q. T. Sun* 27
Precoder Designs for the Relay-Aided X Channel Without Source CSI http://dx.doi.org/10.1109/TSP.2016.2616329 .... *K. Anand, E. Gunawan, and Y. L. Guan* 41

Compressed and Quantized Correlation Estimators http://dx.doi.org/10.1109/TSP.2016.2597128 ... *A. G. Zebadua, P.-O. Amblard, E. Moisan, and O. J. J. Michel* 56

Perfect Recovery Conditions for Non-negative Sparse Modeling http://dx.doi.org/10.1109/TSP.2016.2613067 ... *Y. Itoh, M. F. Duarte, and M. Parente* 69

Generalized Coprime Sampling of Toeplitz Matrices for Spectrum Estimation http://dx.doi.org/10.1109/TSP.2016.2614799 ... *S. Qin, Y. D. Zhang, M. G. Amin, and A. M. Zoubir* 81

Model-Based Nonuniform Compressive Sampling and Recovery of Natural Images Utilizing a Wavelet-Domain Universal Hidden Markov Model http://dx.doi.org/10.1109/TSP.2016.2614654 ... *B. Shahrasbi and N. Rahnavard* 95

Robust Sparse Recovery in Impulsive Noise via ℓ_p-ℓ_1 Optimization http://dx.doi.org/10.1109/TSP.2016.2598316 ... *F. Wen, P. Liu, Y. Liu, R. C. Qiu, and W. Yu* 105

Broadcast Gossip Ratio Consensus: Asynchronous Distributed Averaging in Strongly Connected Networks http://dx.doi.org/10.1109/TSP.2016.2614790 ... *A. Khosravi and Y. S. Kavian* 119

Learning-Based Distributed Detection-Estimation in Sensor Networks With Unknown Sensor Defects http://dx.doi.org/10.1109/TSP.2016.2613062 ... *Q. Zhou, D. Li, S. Kar, L. M. Huie, H. V. Poor, and S. Cui* 130

Network Newton Distributed Optimization Methods http://dx.doi.org/10.1109/TSP.2016.2617829 ... *A. Mokhtari, Q. Ling, and A. Ribeiro* 146

Optimized Spectrum Permutation for the Multidimensional Sparse FFT http://dx.doi.org/10.1109/TSP.2016.2599483 ... *A. Rauh and G. R. Arce* 162

Closed-Loop Autonomous Pilot and Compressive CSIT Feedback Resource Adaptation in Multi-User FDD Massive MIMO Systems http://dx.doi.org/10.1109/TSP.2016.2616326 ... *A. Liu, F. Zhu, and V. K. N. Lau* 173

A Bayesian Approach for Online Recovery of Streaming Signals From Compressive Measurements http://dx.doi.org/10.1109/TSP.2016.2614489 ... *U. L. Wijewardhana and M. Codreanu* 184

Cooperative Simultaneous Localization and Mapping by Exploiting Multipath Propagation http://dx.doi.org/10.1109/TSP.2016.2616324 ... *H. Naseri and V. Koivunen* 200

Joint Optimization of Power and Data Transfer in Multiuser MIMO Systems http://dx.doi.org/10.1109/TSP.2016.2614794 ... *J. Rubio, A. Pascual-Iserte, D. P. Palomar, and A. Goldsmith* 212

Adaptive Cloud Radio Access Networks: Compression and Optimization http://dx.doi.org/10.1109/TSP.2016.2617826 ... *T. X. Vu, H. D. Nguyen, T. Q. S. Quek, and S. Sun* 228

Distributed Fusion With Multi-Bernoulli Filter Based on Generalized Covariance Intersection http://dx.doi.org/10.1109/TSP.2016.2617825 ... *B. Wang, W. Yi, R. Hoseinnezhad, S. Li, L. Kong, and X. Yang* 242

Learning From Hidden Traits: Joint Factor Analysis and Latent Clustering http://dx.doi.org/10.1109/TSP.2016.2614491 ... *B. Yang, X. Fu, and N. D. Sidiropoulos* 256

IEEE TRANSACTIONS ON

SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed MEDLINE U.S. National Library of Medicine

JANUARY 15, 2017 VOLUME 65 NUMBER 2 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

Autoregressive Moving Average Graph Filtering http://dx.doi.org/10.1109/TSP.2016.2614793 .... *E. Isufi, A. Loukas, A. Simonetto, and G. Leus* 274
Chernoff Test for Strong-or-Weak Radar Models http://dx.doi.org/10.1109/TSP.2016.2616323 .... *M. Franceschetti, S. Marano, and V. Matta* 289
Massive MIMO Channel Subspace Estimation From Low-Dimensional Projections http://dx.doi.org/10.1109/TSP.2016.2616336 .... *S. Haghighatshoar and G. Caire* 303
Persistent Homology Lower Bounds on High-Order Network Distances http://dx.doi.org/10.1109/TSP.2016.2620963 .... *W. Huang and A. Ribeiro* 319

IEEE

Joint BS-User Association, Power Allocation, and User-Side Interference Cancellation in Cell-free Heterogeneous Networks http://dx.doi.org/10.1109/TSP.2016.2620962 .... *A. Liu and V. K. N. Lau* 335

A Tensor-Based Method for Large-Scale Blind Source Separation Using Segmentation http://dx.doi.org/10.1109/TSP.2016.2617858 .... *M. Boussé, O. Debals, and L. De Lathauwer* 346

Maximum-Likelihood Adaptive Filter for Partially Observed Boolean Dynamical Systems http://dx.doi.org/10.1109/TSP.2016.2614798 .... *M. Imani and U. M. Braga-Neto* 359

Sparse Reconstruction Algorithm for Nonhomogeneous Counting Rate Estimation http://dx.doi.org/10.1109/TSP.2016.2620104 .... *T. Trigano, Y. Sepulcre, and Y. Ritov* 372

A Pessimistic Approximation for the Fisher Information Measure http://dx.doi.org/10.1109/TSP.2016.2617824 .... *M. S. Stein and J. A. Nossek* 386

Low-Rank Positive Semidefinite Matrix Recovery From Corrupted Rank-One Measurements http://dx.doi.org/10.1109/TSP.2016.2620109 .... *Y. Li, Y. Sun, and Y. Chi* 397

Extending Classical Multirate Signal Processing Theory to Graphs—Part I: Fundamentals http://dx.doi.org/10.1109/TSP.2016.2617833 .... *O. Teke and P. P. Vaidyanathan* 409

Extending Classical Multirate Signal Processing Theory to Graphs—Part II: M-Channel Filter Banks http://dx.doi.org/10.1109/TSP.2016.2620111 .... *O. Teke and P. P. Vaidyanathan* 423

A Unified Framework for Low Autocorrelation Sequence Design via Majorization–Minimization http://dx.doi.org/10.1109/TSP.2016.2620113 .... *L. Zhao, J. Song, P. Babu, and D. P. Palomar* 438

Mismatched Filter Design and Interference Mitigation for MIMO Radars http://dx.doi.org/10.1109/TSP.2016.2620960 .... *T. Aittomäki and V. Koivunen* 454

Consensus +Innovations Distributed Kalman Filter With Optimized Gains http://dx.doi.org/10.1109/TSP.2016.2617827 .... *S. Das and J. M. F. Moura* 467

Sequential Estimation of Hidden ARMA Processes by Particle Filtering—Part I http://dx.doi.org/10.1109/TSP.2016.2598309 .... *I. Urteaga and P. M. Djurić* 482

Sequential Estimation of Hidden ARMA Processes by Particle Filtering—Part II http://dx.doi.org/10.1109/TSP.2016.2598324 .... *I. Urteaga and P. M. Djurić* 494

Optimally Modulated Illumination for Rapid and Accurate Time Synchronization http://dx.doi.org/10.1109/TSP.2016.2612176 .... *M. Sugimoto, H. Kumaki, T. Akiyama, and H. Hashizume* 505

Multidimensional Harmonic Retrieval via Coupled Canonical Polyadic Decomposition—Part I: Model and Identifiability http://dx.doi.org/10.1109/TSP.2016.2614796 .... *M. Sørensen and L. De Lathauwer* 517

Multidimensional Harmonic Retrieval via Coupled Canonical Polyadic Decomposition—Part II: Algorithm and Multirate Sampling http://dx.doi.org/10.1109/TSP.2016.2614797 .... *M. Sørensen and L. De Lathauwer* 528

SignalProcessing | Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

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 HAMZAOUI, *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 CHAKARESKI, *University of Alabama, AL, USA*

Local Arrangements Chair:
Nicholas GARDNER, *University of Bedfordshire, UK*

IEEE/ACM TRANSACTIONS ON

AUDIO, SPEECH, AND LANGUAGE PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

Association for Computing Machinery

www.signalprocessingsociety.org

Indexed in PubMed® and MEDLINE®, products of the United States National Library of Medicine

PubMed

MEDLINE U.S. National Library of Medicine

DECEMBER 2016 VOLUME 24 NUMBER 12 ITASFA (ISSN 2329-9290)

REGULAR PAPERS

Context-Dependent Piano Music Transcription With Convolutional Sparse Coding http://dx.doi.org/10.1109/TASLP.2016.2598305 .... *A. Cogliati, Z. Duan, and B. Wohlberg* 2218
Neural Network Based Multi-Factor Aware Joint Training for Robust Speech Recognition http://dx.doi.org/10.1109/TASLP.2016.2598308 .. *Y. Qian, T. Tan, and D. Yu* 2231
Factorized Hidden Layer Adaptation for Deep Neural Network Based Acoustic Modeling http://dx.doi.org/10.1109/TASLP.2016.2601146 ... *L. Samarakoon and K. C. Sim* 2241
On MMSE-Based Estimation of Amplitude and Complex Speech Spectral Coefficients Under Phase-Uncertainty http://dx.doi.org/10.1109/TASLP.2016.2602549 .... *M. Krawczyk-Becker and T. Gerkmann* 2251
Very Deep Convolutional Neural Networks for Noise Robust Speech Recognition http://dx.doi.org/10.1109/TASLP.2016.2602884 .... *Y. Qian, M. Bi, T. Tan, and K. Yu* 2263
Generation of Affective Accompaniment in Accordance With Emotion Flow http://dx.doi.org/10.1109/TASLP.2016.2603006 .... *Y.-C. Wu and H. H. Chen* 2277

IEEE

Scalable Audio Coding Using Trellis-Based Optimized Joint Entropy Coding and Quantization http://dx.doi.org/10.1109/TASLP.2016.2607339 .... M. Movassagh and P. Kabal 2288
Composition of Deep and Spiking Neural Networks for Very Low Bit Rate Speech Coding http://dx.doi.org/10.1109/TASLP.2016.2604566 .... M. Cernak, A. Lazaridis, A. Asaei, and P. N. Garner 2301
Kernel Method for Voice Activity Detection in the Presence of Transients http://dx.doi.org/10.1109/TASLP.2016.2566919 .... D. Dov, R. Talmon, and I. Cohen 2313
Bayesian Networks to Model the Variability of Speaker Verification Scores in Adverse Environments http://dx.doi.org/10.1109/TASLP.2016.2607343 .... J. Villalba, A. Miguel, A. Ortega, and E. Lleida 2327
Novel Unsupervised Auditory Filterbank Learning Using Convolutional RBM for Speech Recognition http://dx.doi.org/10.1109/TASLP.2016.2607341 .... H. B. Sailor and H. A. Patil 2341
Instantaneous Fundamental Frequency Estimation With Optimal Segmentation for Nonstationary Voiced Speech http://dx.doi.org/10.1109/TASLP.2016.2608948 .... S. M. Nørholm, J. R. Jensen, and M. G. Christensen 2354
Robust Variable Step-Size Decorrelation Normalized Least-Mean-Square Algorithm and its Application to Acoustic Echo Cancellation http://dx.doi.org/10.1109/TASLP.2016.2556280 .... S. Zhang, J. Zhang, and H. Han 2368
Blind Separation of Audio Mixtures Through Nonnegative Tensor Factorization of Modulation Spectrograms http://dx.doi.org/10.1109/TASLP.2016.2602546 .... T. Barker and T. Virtanen 2377
Adaptive Compensation of Misequalization in Narrowband Active Noise Equalizer Systems http://dx.doi.org/10.1109/TASLP.2016.2604212 .... J. Liu and X. Chen 2390
Estimating Speech Recognition Accuracy Based on Error Type Classification http://dx.doi.org/10.1109/TASLP.2016.2603599 .... A. Ogawa, T. Hori, and A. Nakamura 2400
Score-Aging Calibration for Speaker Verification http://dx.doi.org/10.1109/TASLP.2016.2602542 .... F. Kelly and J. H. L. Hansen 2414
An Approach to Score Following for Piano Performances With the Sustained Effect http://dx.doi.org/10.1109/TASLP.2016.2611938 .... B. Li and Z. Duan 2425
Integration of Optimized Modulation Filter Sets Into Deep Neural Networks for Automatic Speech Recognition http://dx.doi.org/10.1109/TASLP.2016.2615239 .... N. Moritz, B. Kollmeier, and J. Anemüller 2439
Multichannel Audio Source Separation With Probabilistic Reverberation Priors http://dx.doi.org/10.1109/TASLP.2016.2614140 .... S. Leglaive, R. Badeau, and G. Richard 2453
Single Snapshot Detection and Estimation of Reflections From Room Impulse Responses in the Spherical Harmonic Domain http://dx.doi.org/10.1109/TASLP.2016.2615238 .... S. Tervo 2466
Extraction of Acoustic Sources Through the Processing of Sound Field Maps in the Ray Space http://dx.doi.org/10.1109/TASLP.2016.2615242 .... D. Marković, F. Antonacci, L. Bianchi, S. Tubaro, and A. Sarti 2481
Long-Term SNR Estimation of Speech Signals in Known and Unknown Channel Conditions http://dx.doi.org/10.1109/TASLP.2016.2615240 .... P. Papadopoulos, A. Tsiartas, and S. Narayanan 2495
Sensitivity of Source–Filter Interaction to Specific Vocal Tract Shapes http://dx.doi.org/10.1109/TASLP.2016.2616543 .... I. R. Titze and A. Palaparthi 2507
A Hybrid Approach for Speech Enhancement Using MoG Model and Neural Network Phoneme Classifier http://dx.doi.org/10.1109/TASLP.2016.2618007 .... S. E. Chazan, J. Goldberger, and S. Gannot 2516
Superdirective Beamforming Based on the Krylov Matrix http://dx.doi.org/10.1109/TASLP.2016.2618003 .... G. Huang, J. Benesty, and J. Chen 2531

Reviewers List http://dx.doi.org/10.1109/TASLP.2016.2618499 2531

EDICS—Editor's Information Classification Scheme http://dx.doi.org/10.1109/TASLP.2016.2625091 2549
Information for Authors http://dx.doi.org/10.1109/TASLP.2016.2625087 2551

Introducing IEEE Collabratec http://dx.doi.org/10.1109/TASLP.2016.2630135 2553

2016 INDEX http://dx.doi.org/10.1109/TASLP.2016.2628026 2554

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

DECEMBER 2016 VOLUME 25 NUMBER 12 IIPRE4 (ISSN 1057-7149)

PAPERS

Dynamic Parallel and Distributed Graph Cuts http://dx.doi.org/10.1109/TIP.2016.2609819 .... *M. Yu, S. Shen, and Z. Hu* 5511
Multi-View 3D Object Retrieval With Deep Embedding Network http://dx.doi.org/10.1109/TIP.2016.2609814 .... *H. Guo, J. Wang, Y. Gao, J. Li, and H. Lu* 5526
Pedestrian Detection Inspired by Appearance Constancy and Shape Symmetry http://dx.doi.org/10.1109/TIP.2016.2609807 .... *J. Cao, Y. Pang, and X. Li* 5538
Unsupervised Domain Adaptation With Label and Structural Consistency http://dx.doi.org/10.1109/TIP.2016.2609820 .... *C.-A. Hou, Y.-H. H. Tsai, Y.-R. Yeh, and Y.-C. F. Wang* 5552
Incorporating Spatial Information and Endmember Variability Into Unmixing Analyses to Improve Abundance Estimates http://dx.doi.org/10.1109/TIP.2016.2601269 .... *T. Uezato, R. J. Murphy, A. Melkumyan, and A. Chlingaryan* 5563
Deep Transfer Metric Learning http://dx.doi.org/10.1109/TIP.2016.2612827 .... *J. Hu, J. Lu, Y.-P. Tan, and J. Zhou* 5576
A Faster, Unbiased Path Opening by Upper Skeletonization and Weighted Adjacency Graphs http://dx.doi.org/10.1109/TIP.2016.2609805 .... *T. Asplund and C. L. Luengo Hendriks* 5589
Hash-Based Line-by-Line Template Matching for Lossless Screen Image Coding http://dx.doi.org/10.1109/TIP.2016.2612884 .... *X. Peng and J. Xu* 5601
A Fast Optimization Method for General Binary Code Learning http://dx.doi.org/10.1109/TIP.2016.2612883 .... *F. Shen, X. Zhou, Y. Yang, J. Song, H. T. Shen, and D. Tao* 5610
Contour Restoration of Text Components for Recognition in Video/Scene Images http://dx.doi.org/10.1109/TIP.2016.2607426 .... *Y. Wu, P. Shivakumara, T. Lu, C. L. Tan, M. Blumenstein, and G. H. Kumar* 5622
Structure Integral Transform Versus Radon Transform: A 2D Mathematical Tool for Invariant Shape Recognition http://dx.doi.org/10.1109/TIP.2016.2609816 .... *B. Wang and Y. Gao* 5635
Hyperspectral Image Recovery via Hybrid Regularization http://dx.doi.org/10.1109/TIP.2016.2614131 .... *R. Arablouei and F. de Hoog* 5649
Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior http://dx.doi.org/10.1109/TIP.2016.2612882 .... *C.-Y. Li, J.-C. Guo, R.-M. Cong, Y.-W. Pang, and B. Wang* 5664
Beyond Object Proposals: Random Crop Pooling for Multi-Label Image Recognition http://dx.doi.org/10.1109/TIP.2016.2612829 .... *M. Wang, C. Luo, R. Hong, J. Tang, and J. Feng* 5678
Web Video Event Recognition by Semantic Analysis From Ubiquitous Documents http://dx.doi.org/10.1109/TIP.2016.2614136 .... *L. Yu, Y. Yang, Z. Huang, P. Wang, J. Song, and H. T. Shen* 5689

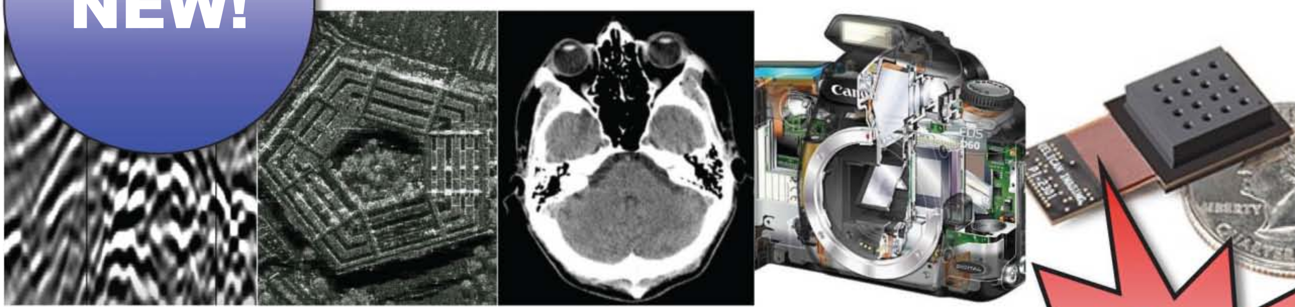
Blind Bleed-Through Removal for Scanned Historical Document Image With Conditional Random Fields http://dx.doi.org/10.1109/TIP.2016.2614133 .... *B. Sun, S. Li, X.-P. Zhang, and J. Sun* 5702
Robust Nanoparticles Detection From Noisy Background by Fusing Complementary Image Information http://dx.doi.org/10.1109/TIP.2016.2614127 .... *Y. Qian, J. Z. Huang, X. Li, and Y. Ding* 5713
Joint Facial Action Unit Detection and Feature Fusion: A Multi-Conditional Learning Approach http://dx.doi.org/10.1109/TIP.2016.2615288 .... *S. Eleftheriadis, O. Rudovic, and M. Pantic* 5727
Learning Collaborative Sparse Representation for Grayscale-Thermal Tracking http://dx.doi.org/10.1109/TIP.2016.2614135 .... *C. Li, H. Cheng, S. Hu, X. Liu, J. Tang, and L. Lin* 5743
Tree-Structured Nuclear Norm Approximation With Applications to Robust Face Recognition http://dx.doi.org/10.1109/TIP.2016.2612885 .... *L. Luo, L. Chen, J. Yang, J. Qian, and B. Zhang* 5757
Low-Rankness Transfer for Realistic Denoising http://dx.doi.org/10.1109/TIP.2016.2612820 .... *H. Badri, H. Yahia, and D. Aboutajdine* 5768
A Coupled Hidden Conditional Random Field Model for Simultaneous Face Clustering and Naming in Videos http://dx.doi.org/10.1109/TIP.2016.2601491 .... *Y. Zhang, Z. Tang, B. Wu, Q. Ji, and H. Lu* 5780
Image Denoising via Bandwise Adaptive Modeling and Regularization Exploiting Nonlocal Similarity http://dx.doi.org/10.1109/TIP.2016.2614160 .... *R. Xiong, H. Liu, X. Zhang, J. Zhang, S. Ma, F. Wu, and W. Gao* 5793
Codewords Distribution-Based Optimal Combination of Equal-Average Equal-Variance Equal-Norm Nearest Neighbor Fast Search Algorithm for Vector Quantization Encoding http://dx.doi.org/10.1109/TIP.2016.2615292 .... *Y.-F. Xie, J.-H. Liu, C.-F. Zhang, L.-S. Kong, and J.-L. Yi* 5806
Multi-View Object Retrieval via Multi-Scale Topic Models http://dx.doi.org/10.1109/TIP.2016.2614132 .... *R. Hong, Z. Hu, R. Wang, M. Wang, and D. Tao* 5814
Multimodal Stereoscopic Movie Summarization Conforming to Narrative Characteristics http://dx.doi.org/10.1109/TIP.2016.2615289 .... *I. Mademlis, A. Tefas, N. Nikolaidis, and I. Pitas* 5828
Subpixel Image Quality Assessment Syncretizing Local Subpixel and Global Pixel Features http://dx.doi.org/10.1109/TIP.2016.2615429 .... *J. Zeng, L. Fang, J. Pang, H. Li, and F. Wu* 5841
Dynamic Programming Using Polar Variance for Image Segmentation http://dx.doi.org/10.1109/TIP.2016.2615809 .... *J. A. Rosado-Toro, M. I. Altbach, and J. J. Rodríguez* 5857
Visual Tracking Under Motion Blur http://dx.doi.org/10.1109/TIP.2016.2615812 .... *B. Ma, L. Huang, J. Shen, L. Shao, M.-H. Yang, and F. Porikli* 5867
High-Efficiency 3D Depth Coding Based on Perceptual Quality of Synthesized Video http://dx.doi.org/10.1109/TIP.2016.2615290 .... *Y. Zhang, X. Yang, X. Liu, Y. Zhang, G. Jiang, and S. Kwong* 5877
Deep Learning Driven Visual Path Prediction From a Single Image http://dx.doi.org/10.1109/TIP.2016.2613686 .... *S. Huang, X. Li, Z. Zhang, Z. He, F. Wu, W. Liu, J. Tang, and Y. Zhuang* 5892
Spatial Pyramid Covariance-Based Compact Video Code for Robust Face Retrieval in TV-Series http://dx.doi.org/10.1109/TIP.2016.2616297 .... *Y. Li, R. Wang, Z. Cui, S. Shan, and X. Chen* 5905
Learning Expressionlets via Universal Manifold Model for Dynamic Facial Expression Recognition http://dx.doi.org/10.1109/TIP.2016.2615424 .... *M. Liu, S. Shan, R. Wang, and X. Chen* 5920
Real-Time Superpixel Segmentation by DBSCAN Clustering Algorithm http://dx.doi.org/10.1109/TIP.2016.2616302 .... *J. Shen, X. Hao, Z. Liang, Y. Liu, W. Wang, and L. Shao* 5933
Defocus Map Estimation From a Single Image Based on Two-Parameter Defocus Model http://dx.doi.org/10.1109/TIP.2016.2617460 .... *S. Liu, F. Zhou, and Q. Liao* 5943
Efficient Non-Consecutive Feature Tracking for Robust Structure-From-Motion http://dx.doi.org/10.1109/TIP.2016.2607425 .... *G. Zhang, H. Liu, Z. Dong, J. Jia, T.-T. Wong, and H. Bao* 5957
Adaptive Subspace-Based Inverse Projections via Division Into Multiple Sub-Problems for Missing Image Data Restoration http://dx.doi.org/10.1109/TIP.2016.2616286 .... *T. Ogawa and M. Haseyama* 5971
A Spatial Compositional Model for Linear Unmixing and Endmember Uncertainty Estimation http://dx.doi.org/10.1109/TIP.2016.2618002 .... *Y. Zhou, A. Rangarajan, and P. D. Gader* 5987

EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIP.2016.2625027 6003
Information for Authors http://dx.doi.org/10.1109/TIP.2016.2625026 6004

List of Reviewers http://dx.doi.org/10.1109/TIP.2016.2617578 .... *Available online at http://ieeexplore.ieee.org*
2016 Index .... *Available online at http://ieeexplore.ieee.org*

NEW!

IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING



30% FASTER TURN AROUND THAN TIP!

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:

Computational Imaging Methods and Models
- Coded image sensing
- Compressed sensing
- Sparse and low-rank models
- Learning-based models, dictionary methods
- Graphical image models
- Perceptual models

Computational Image Formation
- Sparsity-based reconstruction
- Statistically-based inversion methods
- Multi-image and sensor fusion
- Optimization-based methods; proximal iterative methods, ADMM

Computational Photography
- Non-classical image capture
- Generalized illumination
- Time-of-flight imaging
- High dynamic range imaging
- Plenoptic imaging

Computational Consumer Imaging
- Mobile imaging, cell phone imaging
- Camera-array systems
- Depth cameras, multi-focus imaging
- Pervasive imaging, camera networks

Computational Acoustic Imaging
- Multi-static ultrasound imaging
- Photo-acoustic imaging
- Acoustic tomography

Computational Microscopy
- Holographic microscopy
- Quantitative phase imaging
- Multi-illumination microscopy
- Lensless microscopy
- Light field microscopy

Imaging Hardware and Software
- Embedded computing systems
- Big data computational imaging
- Integrated hardware/digital design

Tomographic Imaging
- X-ray CT
- PET
- SPECT

Magnetic Resonance Imaging
- Diffusion tensor imaging
- Fast acquisition

Radar Imaging
- Synthetic aperture imaging
- Inverse synthetic aperture imaging

Geophysical Imaging
- Multi-spectral imaging
- Ground penetrating radar
- Seismic tomography

Multi-spectral Imaging
- 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/*

NOW WITH NO PAGE CHARGES!

IEEE Signal Processing Society ® EMB GRSS IEEE

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 ....

IEEE Signal Processing Society

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

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.signalprocessingsociety.org

DECEMBER 2016 VOLUME 11 NUMBER 12 ITIFA6 (ISSN 1556-6013)

REGULAR PAPERS

A High-Security EEG-Based Login System with RSVP Stimuli and Dry Electrodes http://dx.doi.org/10.1109/TIFS.2016.2577551 ..... *Y. Chen, A. D. Atnafu, I. Schlattner, W. T. Weldtsadik, M.-C. Roh, H. J. Kim, S.-W. Lee, B. Blankertz, and S. Fazli* 2635

Reconstruction Attacks Against Mobile-Based Continuous Authentication Systems in the Cloud http://dx.doi.org/10.1109/TIFS.2016.2594132 ..... *M. Al-Rubaie and J. M. Chang* 2648

Quaternion-Based Image Hashing for Adaptive Tampering Localization http://dx.doi.org/10.1109/TIFS.2016.2594136 ..... *C.-P. Yan, C.-M. Pun, and X.-C. Yuan* 2664

A Data-Centric Approach to Quality Estimation of Role Mining Results http://dx.doi.org/10.1109/TIFS.2016.2594137 ..... *L. Dong, K. Wu, and G. Tang* 2678

Exploiting Multi-Antenna Non-Reciprocal Channels for Shared Secret Key Generation http://dx.doi.org/10.1109/TIFS.2016.2594143 ..... *D. Qin and Z. Ding* 2693

Toward Efficient Multi-Keyword Fuzzy Search Over Encrypted Outsourced Data With Accuracy Improvement http://dx.doi.org/10.1109/TIFS.2016.2596138 ..... *Z. Fu, X. Wu, C. Guan, X. Sun, and K. Ren* 2706

Detecting Byzantine Attacks Without Clean Reference http://dx.doi.org/10.1109/TIFS.2016.2596140 ..... *R. Cao, T. F. Wong, T. Lv, H. Gao, and S. Yang* 2717

Full Frame Encryption and Modulation Obfuscation Using Channel-Independent Preamble Identifier http://dx.doi.org/10.1109/TIFS.2016.2582560 ..... *H. Rahbari and M. Krunz* 2732

SARRE: Semantics-Aware Rule Recommendation and Enforcement for Event Paths on Android http://dx.doi.org/10.1109/TIFS.2016.2596141 ..... *Y. Li, F. Yao, T. Lan, and G. Venkataramani* 2748

Deceptive Deletion Triggers Under Coercion http://dx.doi.org/10.1109/TIFS.2016.2598523 ..... *L. Zhao and M. Mannan* 2763

New Framework for Reversible Data Hiding in Encrypted Domain http://dx.doi.org/10.1109/TIFS.2016.2598528 ..... *F. Huang, J. Huang, and Y.-Q. Shi* 2777

Opting Out of Incentive Mechanisms: A Study of Security as a Non-Excludable Public Good http://dx.doi.org/10.1109/TIFS.2016.2599005 ..... *P. Naghizadeh and M. Liu* 2790

Analytic and Simulation Results about a Compact, Reliable, and Unbiased 1-bit Physically Unclonable Constant http://dx.doi.org/10.1109/TIFS.2016.2599008 ..... *R. Bernardini and R. Rinaldo* 2804

IEEE

Continuous Authentication Using One-Dimensional Multi-Resolution Local Binary Patterns (1DMRLBP) in ECG Biometrics http://dx.doi.org/10.1109/TIFS.2016.2599270 .... *W. Louis, M. Komeili, and D. Hatzinakos* 2818
Server-Aided Public Key Encryption With Keyword Search http://dx.doi.org/10.1109/TIFS.2016.2599293 .... *R. Chen, Y. Mu, G. Yang, F. Guo, X. Huang, X. Wang, and Y. Wang* 2833
Reconstructing High-Resolution Face Models From Kinect Depth Sequences http://dx.doi.org/10.1109/TIFS.2016.2601059 .... *E. Bondi, P. Pala, S. Berretti, and A. Del Bimbo* 2843
On the Race of Worms and Patches: Modeling the Spread of Information in Wireless Sensor Networks http://dx.doi.org/10.1109/TIFS.2016.2594130 .... *M. Sayad Haghighi, S. Wen, Y. Xiang, B. Quinn, and W. Zhou* 2854
Still-to-Video Face Matching Using Multiple Geodesic Flows http://dx.doi.org/10.1109/TIFS.2016.2601060 .... *Y. Zhu, Y. Li, G. Mu, S. Shan, and G. Guo* 2866

List of Reviewers http://dx.doi.org/10.1109/TIFS.2016.2617259 2876

EDICS-Editor's Information Classification Scheme http://dx.doi.org/10.1109/TIFS.2016.2625404 2887
Information for Authors http://dx.doi.org/10.1109/TIFS.2016.2625405 2888

ANNOUNCEMENTS

Call for Papers-IEEE Journal of Selected Topics in Signal Processing Special Issue on Light Field Image Processing http://dx.doi.org/10.1109/TIFS.2016.2626022 2890
Call for Papers-IEEE Transactions on Audio, Speech, and Language Processing Special Issue on Biosignal-Based Spoken Communication http://dx.doi.org/10.1109/TIFS.2016.2626002 2891
Call for Papers-IEEE Transactions on Signal and Information Processing Over Networks Special Issue on Distributed Signal Processing for Security and Privacy in Networked Cyber-Physical Systems http://dx.doi.org/10.1109/TIFS.2016.2626003 2892

2016 Index http://dx.doi.org/10.1109/TIFS.2016.2637459 .... *Available online at http://ieeexplore.ieee.org*

3rd IEEE International Conference on Network Softwarization

SOFTWARIZATION SUSTAINING A HYPER-CONNECTED WORLD: EN ROUTE TO 5G

3-7 JULY 2017, BOLOGNA, ITALY

http://sites.ieee.org/netsoft—CALL FOR PAPERS

IEEE NETSOFT

IEEE Advancing Technology for Humanity

The 3rd IEEE International Conference on Network Softwarization (NetSoft 2017) will be held July 3–7, 2017 at the University of Bologna in Bologna, Italy. IEEE NetSoft is the flagship conference of the IEEE SDN Initiative which aims to address "Softwarization" of networks and systemic trends concerning the convergence of Cloud Computing, Software-Defined Networks, and Network Function Virtualization.

TOPICS OF INTEREST

Authors are invited to submit papers that fall into the area of software-defined and virtualized infrastructures. Topics of interest include, but are not limited to, the following:

- SDN and NFV as enabling technologies for 5G
- From Cloud Computing to Edge-Fog Computing
- 5G Functional Decomposition and Infrastructure slicing
- 5G sustainable ecosystems: IoT, Industry 4.0, Pervasive Robotics, Self-driving vehicles, Tactile Internet, Immersive Communications, Artificial Intelligence applications
- Software Defined infrastructures for Public Protection and Disaster Relief (PPDR) network services
- Service Function Chaining for NFV: Modeling, composition algorithms, deployment
- Intent-based interfacing for NFV
- SDN/NFV Network & Service Orchestration and Management
- Management of federated SDN/NFV infrastructure and frameworks
- Real time operations and efficient network/service monitoring in SDN/NFV
- Performance and scalability issues in NFV implementation scenarios
- Traffic Engineering and QoS/QoE in SDN/NFV
- APIs, protocols and languages for programmable networks and Software-Defined Infrastructure
- SDN switch/router architectures/designs
- SDN/NFV issues and opportunities for security, trust and privacy
- Experience reports from experimental testbeds and deployment
- Softwarized platforms for Internet-of-Things (IoT)
- New value chains and business models

SCOPE

The telecommunications landscape will change radically in the next few years. Pervasive ultra-broadband, programmable networks, and cost reduction of IT systems are paving the way to new services and commoditization of telecommunications infrastructure while lowering entrance barriers for new players and giving rise to new value chains. While this results in considerable challenges for service providers, this transformation also brings unprecedented opportunities for the Digital Society and the Digital Economy related to emerging new services and applications. Examples include Tactile Internet of Things, Industry 4.0, Cloud Robotics, and Artificial Intelligence. 5G will both exploit and accelerate this transformation.

NetSoft 2017 aims to capture the theme of "Softwarization Sustaining a Hyper-connected World: en route to 5G" and serve as forum for researchers to discuss the latest advances in this area. NetSoft 2017 will feature technical paper, keynotes, tutorials, and demos and exhibits from world-leading experts representing operators, vendors, research institutes, open source projects, and academia.

PAPER SUBMISSION

Authors are invited to submit original contributions (written in English) in PDF format. Only original papers not published or submitted for publication elsewhere can be submitted. Papers can be of two types: full (up to 9 pages) or short (up to 5 pages) papers. Full Papers accepted as short Papers will be required to be reduced to 5-pages length. Papers should be in IEEE 2-column US-Letter style using IEEE Conference template (*http://www.ieee.org/conferences_events/conferences/publishing/templates.html*) and submitted in PDF format via JEMS at: *https://jems.sbc.org.br/home.cgi?c=2657*. Papers exceeding these limits, multiple submissions, and self-plagiarized papers will be rejected without further review. All submitted papers will be subject to a peer-review process. The accepted papers will be published in IEEE Xplore, provided that the authors do present their paper at the conference.

IMPORTANT DATES

December 5, 2016: Technical Papers deadline
December 15, 2016: Workshop Submission deadline
March 6, 2017: Paper submission acceptance notification
April 10, 2017: Full Conference Paper; Speaker registration
May 30, 2017: Early-Bird Registration

GENERAL CO-CHAIRS

Antonio Manzalini (Telecom Italia Mobile, Italy)
Roberto Verdone (University of Bologna, Italy)

STEERING COMMITTEE CHAIR

Prosper Chemouil (Orange Labs, France)

TPC CO-CHAIRS

Franco Callegati (University of Bologna, Italy)
Alexander Clemm (Cisco, USA)
Kohei Shiomoto (NTT Labs, Japan)

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/

DECEMBER 2016 VOLUME 18 NUMBER 12 ITMUF8 (ISSN 1520-9210)

PAPERS

Compression and Coding

Quadtree Degeneration for HEVC http://dx.doi.org/10.1109/TMM.2016.2598481 .... *Y. Gao, P. Liu, Y. Wu, and K. Jia* 2321

Clustering-Based Content Adaptive Tiles Under On-chip Memory Constraints http://dx.doi.org/10.1109/TMM.2016.2600439 .... *X. Jin and Q. Dai* 2331

Watermarking, Encryption, and Data Hiding

Data Hiding Robust to Mobile Communication Vocoders http://dx.doi.org/10.1109/TMM.2016.2599149 .... *R. Kazemi, F. Pérez-González, M. A. Akhaee, and F. Behnia* 2345

Image/Video/Graphics Analysis and Synthesis

Interactive Multilabel Image Segmentation via Robust Multilayer Graph Constraints http://dx.doi.org/10.1109/TMM.2016.2600441 .... *T. Wang, Z. Ji, Q. Sun, Q. Chen, and X.-Y. Jing* 2358

Task-Driven Progressive Part Localization for Fine-Grained Object Recognition http://dx.doi.org/10.1109/TMM.2016.2602060 .... *C. Huang, Z. He, G. Cao, and W. Cao* 2372

System Design Methodology and Tools

A Constellation Design Methodology Based on QoS and User Demand in High-Altitude Platform Broadband Networks http://dx.doi.org/10.1109/TMM.2016.2595260 .... *F. Dong, H. Han, X. Gong, J. Wang, and H. Li* 2384

Video Surveillance and Semantic Analysis

Exemplar-AMMs: Recognizing Crowd Movements From Pedestrian Trajectories http://dx.doi.org/10.1109/TMM.2016.2598091 .... *W. Liu, R. W. H. Lau, X. Wang, and D. Manocha* 2398

Multimedia Using Haptic and Physiological Information

Deep Learning for Surface Material Classification Using Haptic and Visual Information http://dx.doi.org/10.1109/TMM.2016.2598140 .... *H. Zheng, L. Fang, M. Ji, M. Strese, Y. Özer, and E. Steinbach* 2407

Multimodal Perception, Integration, and Multisensory Fusion

Mean-Shift and Sparse Sampling-Based SMC-PHD Filtering for Audio Informed Visual Speaker Tracking http://dx.doi.org/10.1109/TMM.2016.2599150 .... *V. Kiliç, M. Barnard, W. Wang, A. Hilton, and J. Kittler* 2417

Media Translation for Sense Substitution

Perceiving Graphical and Pictorial Information via Hearing and Touch http://dx.doi.org/10.1109/TMM.2016.2601029 .... *P. M. Silva, T. N. Pappas, J. Atkins, and J. E. West* 2432

Subjective and Objective Quality Assessment, and User Experience

Perceptual Annoyance Models for Videos With Combinations of Spatial and Temporal Artifacts http://dx.doi.org/10.1109/TMM.2016.2601027 .... *A. F. Silva, M. C. Q. Farias, and J. A. Redi* 2446

Blind Image Quality Assessment Using Statistical Structural and Luminance Features http://dx.doi.org/10.1109/TMM.2016.2601028 .... *Q. Li, W. Lin, J. Xu, and Y. Fang* 2457

Ubiquitous Media Access

Optimal Incentive Design for Cloud-Enabled Multimedia Crowdsourcing http://dx.doi.org/10.1109/TMM.2016.2604080 .... *S. Maharjan, Y. Zhang, and S. Gjessing* 2470

Multimedia Search and Retrieval

Robust Latent Poisson Deconvolution From Multiple Features for Web Topic Detection http://dx.doi.org/10.1109/TMM.2016.2598439 .... *J. Pang, F. Tao, C. Zhang, W. Zhang, Q. Huang, and B. Yin* 2482

Web and Internet

Image Classification by Cross-Media Active Learning With Privileged Information http://dx.doi.org/10.1109/TMM.2016.2602938 .... *Y. Yan, F. Nie, W. Li, C. Gao, Y. Yang, and D. Xu* 2494

Multimedia Streaming and Transport

Trend-Aware Video Caching Through Online Learning http://dx.doi.org/10.1109/TMM.2016.2596042 .... *S. Li, J. Xu, M. van der Schaar, and W. Li* 2503

Resource-Efficient Mobile Multimedia Streaming With Adaptive Network Selection http://dx.doi.org/10.1109/TMM.2016.2604565 .... *J. Lee, K. Lee, C. Han, T. Kim, and S. Chong* 2517

Multimedia Sentiment Analysis and Synthesis; Affective Multimedia Processing

A Deep Neural Network-Driven Feature Learning Method for Multi-view Facial Expression Recognition http://dx.doi.org/10.1109/TMM.2016.2598092 .... *T. Zhang, W. Zheng, Z. Cui, Y. Zong, J. Yan, and K. Yan* 2528

Multimedia Networking and Processing in the Cloud and Data Centers

Media Query Processing for the Internet-of-Things: Coupling of Device Energy Consumption and Cloud Infrastructure Billing http://dx.doi.org/10.1109/TMM.2016.2600438 .... *F. Renna, J. Doyle, V. Giotsas, and Y. Andreopoulos* 2537

Ultra-Efficient Surveillance Video and Coding Of Multimedia Features

Person Reidentification via Ranking Aggregation of Similarity Pulling and Dissimilarity Pushing http://dx.doi.org/10.1109/TMM.2016.2605058 .... *M. Ye, C. Liang, Y. Yu, Z. Wang, Q. Leng, C. Xiao, J. Chen, and R. Hu* 2553

ANNOUNCEMENTS

Introducing IEEE Collabratec http://dx.doi.org/10.1109/TMM.20162.2625923 2567

CALLS FOR PAPERS

IEEE International Conference on Multimedia and Expo Call for Papers http://dx.doi.org/10.1109/TMM.2016.2625925 2568

IEEE TRANSACTIONS ON MULTIMEDIA Call for Papers for the Special Issue on Video Over Future Networks: Emerging Technologies, Infrastructures, and Applications http://dx.doi.org/10.1109/TMM.2016.2625899 2569

Information for Authors http://dx.doi.org/10.1109/TMM.2016.2625921 2570

2016 INDEX .... Available online at http://ieeexplore.ieee.org

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
IEEE Signal Processing Society

IEEE JOURNAL OF
SELECTED TOPICS IN SIGNAL PROCESSING

IEEE Signal Processing Society®

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/JSTSP.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/JSTSP.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/JSTSP.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/JSTSP.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/JSTSP.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/JSTSP.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/JSTSP.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 .... K. Cumanan, Z. Ding, M. Xu, and H. V. Poor 1417
Improving Wireless Physical Layer Security via Exploiting Co-Channel Interference http://dx.doi.org/10.1109/JSTSP.2016.2600516 .... L. Li, A. P. Petropulu, Z. Chen, and J. Fang 1433
Wireless Information Surveillance via Proactive Eavesdropping with Spoofing Relay http://dx.doi.org/10.1109/JSTSP.2016.2600519 .... Y. Zeng and R. Zhang 1449
Energy Efficiency of Confidential Multi-Antenna Systems With Artificial Noise and Statistical CSI http://dx.doi.org/10.1109/JSTSP.2016.2607690 .... A. Zappone, P.-H. Lin, and E. Jorswieck 1462
Directional Modulation Via Symbol-Level Precoding: A Way to Enhance Security http://dx.doi.org/10.1109/JSTSP.2016.2600521 .... A. Kalantari, M. Soltanalian, S. Maleki, S. Chatzinotas, and B. Ottersten 1478
Secure Multiple Amplify-and-Forward Relaying With Cochannel Interference http://dx.doi.org/10.1109/JSTSP.2016.2607692 .... L. Fan, X. Lei, N. Yang, T. Q. Duong, and G. K. Karagiannidis 1494
Multipair Two-Way Relay Network With Harvest-Then-Transmit Users: Resolving Pairwise Uplink-Downlink Coupling http://dx.doi.org/10.1109/JSTSP.2016.2612162 .... S. Wang, M. Xia, and Y.-C. Wu 1506
On the Performance of mmWave Networks Aided by Wirelessly Powered Relays http://dx.doi.org/10.1109/JSTSP.2016.2610944 .... S. Biswas, S. Vuppala, and T. Ratnarajah 1522
Accumulate and Jam: Towards Secure Communication via A Wireless-Powered Full-Duplex Jammer http://dx.doi.org/10.1109/JSTSP.2016.2600523 .... Y. Bi and H. Chen 1538
Quadrature Amplitude Modulation Division for Multiuser MISO Broadcast Channels http://dx.doi.org/10.1109/JSTSP.2016.2607684 .... Z. Dong, Y.-Y. Zhang, J.-K. Zhang, and X.-C. Gao 1551

CORRECTION

Corrections to “A Characterization of Deterministic Sampling Patterns for Low-Rank Matrix Completion” http://dx.doi.org/10.1109/JSTSP.2016.2618502 .... D. L. Pimentel-Alarcón, N. Boston, and R. D. Nowak 1567

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

IEEE

SIGNAL PROCESSING LETTERS

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.ieee.org/sp/index.html

DECEMBER 2016 VOLUME 23 NUMBER 12 ISPLEM (ISSN 1070-9908)

LETTERS

Improved Wideband DOA Estimation Using Modified TOPS (mTOPS) Algorithm http://dx.doi.org/10.1109/LSP.2016.2614310 .... *A. K. Shaw* 1697
Structure-Aware Slow Feature Analysis for Age Estimation http://dx.doi.org/10.1109/LSP.2016.2602538 .... *Z. He, X. Li, Z. Zhang, Y. Zhang, J. Xiao, and X. Zhou* 1702
CRLB for I/Q Imbalance Estimation in FMCW Radar Receivers http://dx.doi.org/10.1109/LSP.2016.2614982 .... *F. Bandiera, A. Coluccia, V. Dodde, A. Masciullo, and G. Ricci* 1707
Distributed Differential Modulation Over Asymmetric Fading Channels http://dx.doi.org/10.1109/LSP.2016.2609418 .... *S. AlMaeeni, P. C. Sofotasios, S. Muhaidat, G. K. Karagiannidis, and M. Valkama* 1712
Color Gaussian Jet Features For No-Reference Quality Assessment of Multiply-Distorted Images http://dx.doi.org/10.1109/LSP.2016.2617743 .... *H. Hadizadeh and I. V. Bajić* 1717
RGBD Co-saliency Detection via Bagging-Based Clustering http://dx.doi.org/10.1109/LSP.2016.2615293 .... *H. Song, Z. Liu, Y. Xie, L. Wu, and M. Huang* 1722
Constrained Perturbation Regularization Approach for Signal Estimation Using Random Matrix Theory http://dx.doi.org/10.1109/LSP.2016.2615683 .... *M. Suliman, T. Ballal, A. Kammoun, and T. Y. Al-Naffouri* 1727
Antenna Placement for MIMO Localization Systems With Varying Quality of Receiver Hardware Elements http://dx.doi.org/10.1109/LSP.2016.2616648 .... *V. Aggarwal and L. Huie* 1732
Visual Saliency Detection via Sparse Residual and Outlier Detection http://dx.doi.org/10.1109/LSP.2016.2617340 .... *H. Tang, C. Chen, and X. Pei* 1736
Online Dereverberation for Dynamic Scenarios Using a Kalman Filter With an Autoregressive Model http://dx.doi.org/10.1109/LSP.2016.2616888 .... *S. Braun and E. A. P. Habets* 1741
Data Consistency Conditions for Cone-Beam Projections on a Circular Trajectory http://dx.doi.org/10.1109/LSP.2016.2616026 .... *R. Clackdoyle, L. Desbat, J. Lesaint, and S. Rit* 1746
MIMO Detection and Equalization for Single-Carrier Systems Using the Alternating Direction Method of Multipliers http://dx.doi.org/10.1109/LSP.2016.2618959 .... *N. Souto and R. Dinis* 1751
Performance Analysis of Low-Flux Least-Squares Single-Pixel Imaging http://dx.doi.org/10.1109/LSP.2016.2617329 .... *D. Shin, J. H. Shapiro, and V. K. Goyal* 1756
Mixed Pulse Accumulation for Compressive Sensing Radar http://dx.doi.org/10.1109/LSP.2016.2618958 .... *Y. Tao and G. Zhang* 1761
A Virtual User Pairing Scheme to Optimally Utilize the Spectrum of Unpaired Users in Non-orthogonal Multiple Access http://dx.doi.org/10.1109/LSP.2016.2619371 .... *M. B. Shahab, M. F. Kader, and S. Y. Shin* 1766
A Topological Low-Pass Filter for Quasi-Periodic Signals http://dx.doi.org/10.1109/LSP.2016.2619678 .... *M. Robinson* 1771
Revealing Hidden 3-D Reflection Symmetry http://dx.doi.org/10.1109/LSP.2016.2620380 .... *R. Nagar and S. Raman* 1776

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

Novel Channel Estimation for Non-orthogonal Multiple Access Systems http://dx.doi.org/10.1109/LSP.2016.2617897 .... *Y. Tan, J. Zhou, and J. Qin* 1781
Transient Performance Analysis of Zero-Attracting LMS http://dx.doi.org/10.1109/LSP.2016.2616890 .... *J. Chen, C. Richard, Y. Song, and D. Brie* 1786
Low-Complexity 2D Direction-of-Arrival Estimation for Acoustic Sensor Arrays http://dx.doi.org/10.1109/LSP.2016.2614107 .... *K. Yu, R. E. Hudson, Y. D. Zhang, K. Yao, C. Taylor, and Z. Wang* 1791
The Impact of Efficient Transport Blocks Management on the Downlink Power in MIMO Spatial Multiplexing of LTE-A http://dx.doi.org/10.1109/LSP.2016.2622479 .... *B. Al-Doori and X. Liu* 1796
Blind Picture Upscaling Ratio Prediction http://dx.doi.org/10.1109/LSP.2016.2603842 .... *T. R. Goodall, I. Katsavounidis, Z. Li, A. Aaron, and A. C. Bovik* 1801
Robust Semi-Supervised Classification for Noisy Labels Based on Self-Paced Learning http://dx.doi.org/10.1109/LSP.2016.2619352 .... *N. Gu, M. Fan, and D. Meng* 1806
Spatial Filtering Based on Differential Spectrum for Improving ML DOA Estimation Performance http://dx.doi.org/10.1109/LSP.2016.2605006 .... *R. P. Lemos, H. V. Leao e Silva, E. L. Flores, J. A. Kunzler, and D. F. Burgos* 1811
Optimal Binary Periodic Almost-Complementary Pairs http://dx.doi.org/10.1109/LSP.2016.2600586 .... *A. R. Adhikary, Z. Liu, Y. L. Guan, S. Majhi, and S. Z. Budishin* 1816
Probabilistic Latent Semantic Analysis for Multichannel Biomedical Signal Clustering http://dx.doi.org/10.1109/LSP.2016.2623801 .... *J. Wang and M. She* 1821
Uncertainty Exchange Through Multiple Quadrature Kalman Filtering http://dx.doi.org/10.1109/LSP.2016.2618397 .... *J. Vilà-Valls, P. Closas, and Á. F. García-Fernández* 1825
On the Maximal Invariant Statistic for Adaptive Radar Detection in Partially Homogeneous Disturbance With Persymmetric Covariance http://dx.doi.org/10.1109/LSP.2016.2618619 .... *D. Ciuonzo, D. Orlando, and L. Pallotta* 1830
A Geometrical Look at MOSPA Estimation Using Transportation Theory http://dx.doi.org/10.1109/LSP.2016.2614774 .... *G. M. Lipsa and M. Guerriero* 1835
Large Inpainting of Face Images With Trainlets http://dx.doi.org/10.1109/LSP.2016.2616354 .... *J. Sulam and M. Elad* 1839
Optimal Memory Size Formula for Moving-Average Digital Phase-Locked Loop http://dx.doi.org/10.1109/LSP.2016.2623520 .... *C. K. Ahn, P. Shi, and S. H. You* 1844
Unified Maximum Likelihood Form for Bias Constrained FIR Filters http://dx.doi.org/10.1109/LSP.2016.2627001 .... *S. Zhao and Y. S. Shmaliy* 1848
Symbol Detection for Faster-Than-Nyquist Signaling by Sum-of-Absolute-Values Optimization http://dx.doi.org/10.1109/LSP.2016.2625839 .... *H. Sasahara, K. Hayashi, and M. Nagahara* 1853
Sparse Decomposition for Signal Periodic Model Over Complex Exponential Dictionary http://dx.doi.org/10.1109/LSP.2016.2619329 .... *S. Deng and J. Han* 1858
Low-Rank and Sparsity Analysis Applied to Speech Enhancement Via Online Estimated Dictionary http://dx.doi.org/10.1109/LSP.2016.2627029 .... *P. Sun and J. Qin* 1862
Online Anomaly Detection With Nested Trees http://dx.doi.org/10.1109/LSP.2016.2623773 .... *I. Delibalta, K. Gokcesu, M. Simsek, L. Baruh, and S. S. Kozat* 1867
Second-Order Cone Relaxation for TDOA-Based Localization Under Mixed LOS/NLOS Conditions http://dx.doi.org/10.1109/LSP.2016.2627603 .... *W. Wang, G. Wang, F. Zhang, and Y. Li* 1872
Speech Bandwidth Extension Using Recurrent Temporal Restricted Boltzmann Machines http://dx.doi.org/10.1109/LSP.2016.2621053 .... *Y. Wang, S. Zhao, J. Li, and J. Kuang* 1877
Image Fusion With Convolutional Sparse Representation http://dx.doi.org/10.1109/LSP.2016.2618776 .... *Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang* 1882
Superpixel-Guided Adaptive Image Smoothing http://dx.doi.org/10.1109/LSP.2016.2630741 .... *H. Eun and C. Kim* 1887

List of Reviewers http://dx.doi.org/10.1109/LSP.2016.2617138 1892

EDICS–Editors' Information Classification Scheme http://dx.doi.org/10.1109/LSP.2016.2627696 1902
Information for Authors http://dx.doi.org/10.1109/LSP.2016.2627697 1903

2016 INDEX http://dx.doi.org/10.1109/LSP.2016.2636642 1905

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/TCI.2016.2579601 .... *H. Zhao and M. Comer* 395
Plug-and-Play Priors for Bright Field Electron Tomography and Sparse Interpolation http://dx.doi.org/10.1109/TCI.2016.2599778 ..... *S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman* 408
Mean Squared Error Based Excitation Pattern Design for Parallel Transmit and Receive SENSE MRI Image Reconstruction http://dx.doi.org/10.1109/TCI.2016.2610141 .... *I. Y. Chun, S. Noh, D. J. Love, T. M. Talavage, S. Beckley, and S. J. Kisner* 424
Spectral Super-Resolution in Colored Coded Aperture Spectral Imaging http://dx.doi.org/10.1109/TCI.2016.2612943 .... *A. Parada-Mayorga and G. R. Arce* 440
Robust Bayesian Target Detection Algorithm for Depth Imaging From Sparse Single-Photon Data http://dx.doi.org/10.1109/TCI.2016.2618323 .... *Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin* 456

Computational Image Formation

Combining Local Regularity Estimation and Total Variation Optimization for Scale-Free Texture Segmentation http://dx.doi.org/10.1109/TCI.2016.2594139 .... *N. Pustelnik, H. Wendt, P. Abry, and N. Dobigeon* 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/TCI.2016.2601296 .... *K. H. Jin, D. Lee, and J. C. Ye* 480
Fast Multilayer Laplacian Enhancement http://dx.doi.org/10.1109/TCI.2016.2607142 .... *H. Talebi and P. Milanfar* 496
Spectral CT Reconstruction With Image Sparsity and Spectral Mean http://dx.doi.org/10.1109/TCI.2016.2609414 .... *Y. Zhang, Y. Xi, Q. Yang, W. Cong, J. Zhou, and G. Wang* 510
Two-dimensional Autofocus for Spotlight SAR Polar Format Imagery http://dx.doi.org/10.1109/TCI.2016.2612945 .... *X. Mao and D. Zhu* 524
Numerical Inversion of Circular arc Radon Transform http://dx.doi.org/10.1109/TCI.2016.2615806 .... *T. A. Syed, V. P. Krishnan, and J. Sivaswamy* 540

Computational Imaging Hardware and Software

Vehicle Tracking in Video Using Fractional Feedback Kalman Filter http://dx.doi.org/10.1109/TCI.2016.2600480 .... *H. Kaur and J. S. Sahambi* 550

List of Reviewers http://dx.doi.org/10.1109/TCI.2016.2618098 562

EDICS—Editor's Classification Information Scheme http://dx.doi.org/10.1109/TCI.2016.2622275 564
Information for Authors http://dx.doi.org/10.1109/TCI.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. Mahdizadehaghdam, 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

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

2016 INDEX .... *Available online at http://ieeexplore.ieee.org*

General Chairs
Yasutaka Ogawa
Wei Yu
Fumiyuki Adachi

Technical Program Chairs
Tomoaki Ohtsuki
Lutz Lampe
Wing-Kin (Ken) Ma

Special Session Chair
Tony Q. S. Quek

Technical Program Committee
Waheed Bajwa
Chong-Yung Chi
Philippe Ciblat
Min Dong
FeiFei Gao
Joakim Jaldén
Shi Jin
Yindi Jing
Markku Juntti
Eleftherios Karipidis
Erik G. Larsson
David J. Love
Matthew McKay
Urbashi Mitra
Chandra Murthy
Michael Rabbat
Alejandro Ribeiro
Ahmed Sadek
Mathini Sellathurai
Milica Stojanovic
Weifeng Su
Cihan Tepedelenlioglu
Wolfgang Utschick
Sergiy A. Vorobyov
Pengfei Xia
Rui Zhang
Wei Zhang

SPAWC 2017 will be held at Hokkaido University in Sapporo, Japan on July 3-6, 2017. The workshop is devoted to advances in signal processing for wireless communications, networking, and information theory. The technical program features plenary talks as well as invited and contributed papers presented in poster format.

Call for Papers

Prospective authors are invited to submit papers in the following areas:

- Smart antennas, MIMO systems, massive MIMO, and space-time processing
- Single-carrier, multi-carrier, and multi-rate systems
- Multiple-access and broadcast channels, multi-user receivers
- Signal processing for ad-hoc, multi-hop, and sensor networks
- Cooperative communication, coordinated multipoint transmission and reception
- Distributed resource allocation and scheduling
- Convex and non-convex optimization; Game theory for communications
- Interference management, dynamic spectrum management
- Heterogeneous networks, small cells
- Millimeter wave, 60GHz communications
- Full duplex systems
- Physical layer security
- Feedback in wireless networks
- Cognitive radio and networks
- Cooperative sensing, compressed sensing, sparse signal processing
- Machine-to-machine, device-to-device communications
- Ultra-wideband radio, localization, RFID
- Modeling, estimation and equalization of wireless channels
- Acquisition, synchronization, and tracking
- Cross-layer issues, joint source-channel coding, delay-limited communication
- Signal processing for optical, satellite, and underwater communications
- Signal processing for nano- and molecular communications
- Energy efficiency and energy harvesting
- Large networks and big data in wireless communications
- Emerging techniques, technologies, and new waveforms for 5G

Full papers up to five-page limit should be submitted via EDAS.

March 1, 2017	Paper submission deadline
April 30, 2017	Acceptance notification
May 10, 2017	Final paper due

Finance Chairs
Yasushi Yamao
Hai Lin

Local Arrangement Chairs
Takeo Ohgane
Toshihiko Nishimura

Publication Chairs
Kazunori Hayashi
Kenichi Higuchi

Publicity Chairs
Julian Webber
Koichi Adachi

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page

5th IEEE Global Conference on Signal and Information Processing

November 14 - 16, 2017, Montreal, Canada

http://www.ieeeglobalsip.org
2017.ieeeglobalsip.org

Call for Symposium Proposals

We invite Symposium proposals for the fifth IEEE Global Conference on Signal and Information Processing (GlobalSIP) which will be held in Montreal, Quebec, Canada on November 14-16, 2017. GlobalSIP is a flagship IEEE Signal Processing Society conference. It 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. Topics include but are not limited to:

- Signal and information processing for
  - communications and networks, including green communications
  - optical communications
  - forensics and security
  - finance
  - energy and power systems (e.g., smart grid)
  - genomics and bioengineering (physiological, pharmacological and behavioral)
  - neural networks, including deep learning
- Image and video processing
- Selected topics in speech processing and human language technologies
- Human machine interfaces
- Multimedia transmission, indexing, retrieval, and quality of experience
- Selected topics in statistical signal processing
- Cognitive communications and radar
- Graph-theoretic signal processing
- Machine learning
- Compressed sensing and sparsity aware processing
- Seismic signal processing
- Big data and social media challenges
- Hardware and real-time implementations
- Other (industrial) emerging applications of signal and information processing.

Symposium proposals should contain the following information: title; duration (e.g., full day or half day); paper length, acceptance rate; name, address, and a short CV (up to 250 words) of the organizers, including the technical chairs (if any); a 1-page or 2-page description of the topics to be addressed, including timeliness and relevance to the signal processing community; names of (potential) members of the technical program committee; invited speakers' name; a draft call for papers. Please pack everything together in a single pdf document. More detailed information can be found in GlobalSIP2017 Symposium Proposal Preparation Guide.

Proposed Timeline

- **Jan. 20, 2017**: Symposium proposals due
- **Jan. 25, 2017**: Symposium selection decision made
- **Feb. 1, 2017**: Call for Papers for accepted Symposia
- **May 15, 2017**: Paper submission due
- **June 30, 2017**: Notification of Acceptance
- **July 22, 2017**: Camera-ready paper due.

GlobalSIP 2017 - Call for Symposium Proposals

We invite Symposium proposals for the fifth IEEE Global Conference on Signal and Information Processing (GlobalSIP2017) which will be held in Montreal, Quebec, Canada on **November 14-16, 2017** (http://2017.ieeeglobalsip.org/). GlobalSIP is a flagship IEEE Signal Processing Society conference. It 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. Topics include but are not limited to:

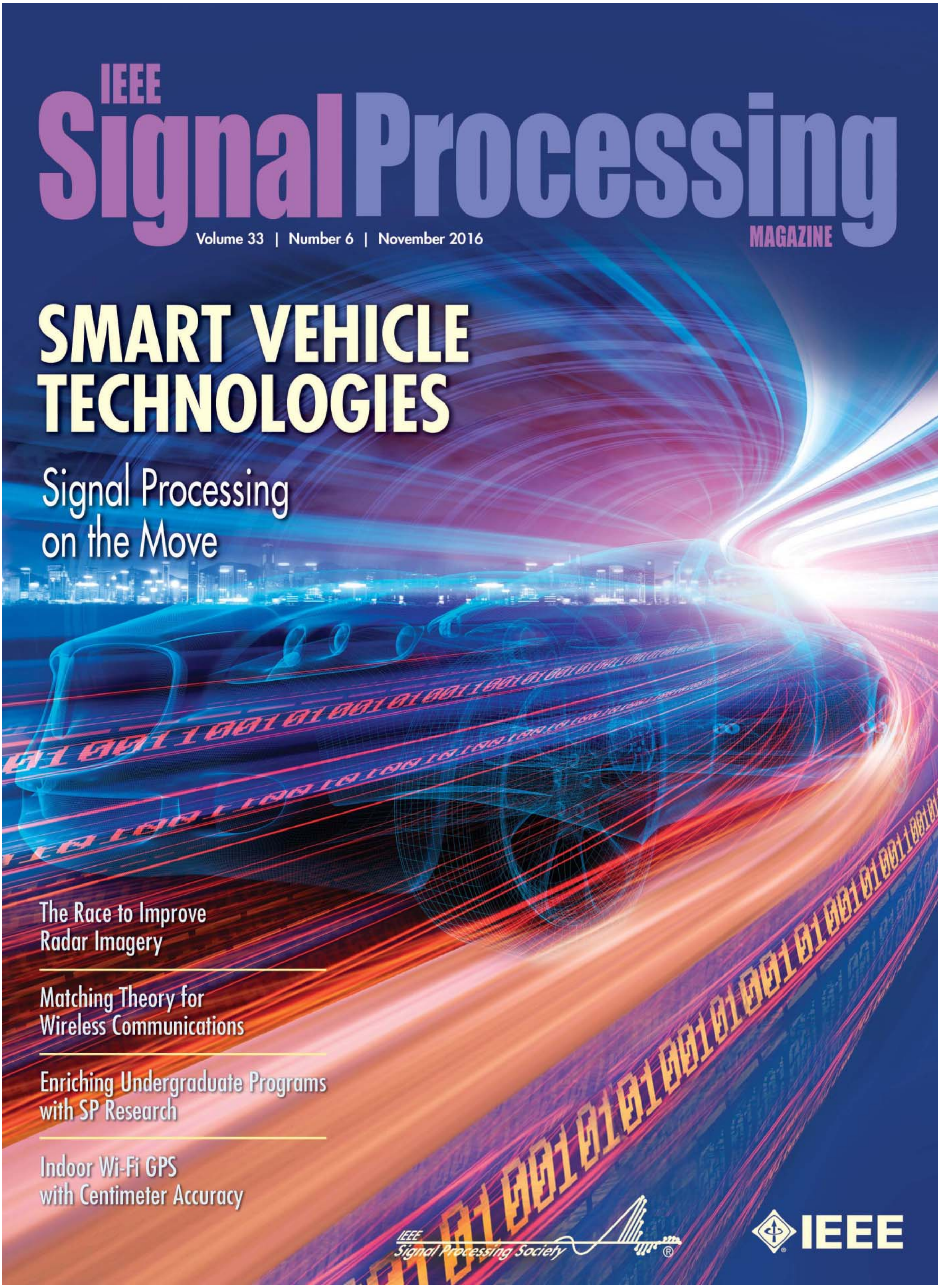
- Signal and information processing for
  - communications and networks, including green communications
  - optical communications
  - forensics and security
  - finance
  - energy and power systems (e.g., smart grid)
  - genomics and bioengineering (physiological, pharmacological and behavioral)
  - neural networks, including deep learning
- Image and video processing
- Selected topics in speech processing and human language technologies
- Human machine interfaces
- Multimedia transmission, indexing, retrieval, and quality of experience
- Selected topics in statistical signal processing
- Cognitive communications and radar
- Graph-theoretic signal processing
- Machine learning
- Compressed sensing and sparsity aware processing
- Seismic signal processing
- Big data and social media challenges
- Hardware and real-time implementations
- Other (industrial) emerging applications of signal and information processing.
- Interdisciplinary theme symposia are strongly encouraged.

Symposium proposals should contain the following information: title; duration (e.g., full day or half day); paper length; name, address, and a short CV (up to 250 words) of the organizers, including the technical chairs (if any); a 1-page or 2-page description of the topics to be addressed, including timeliness and relevance to the signal processing community; names of (potential) members of the technical program committee; invited/potential speakers' names; a draft call for papers (up to 1 page). Please pack everything together in a single pdf document and email your proposal to the Technical Program Committee (TPC) Chairs. More detailed information can be found in "GlobalSIP2017 Symposium Proposal Preparation Guide" at http://2017.ieeeglobalsip.org/. For better consideration, proposers are strongly advised to submit their proposals as early as possible.

Proposed Timeline

- **Jan. 20, 2017**: Symposium proposals due
- **Jan. 25, 2017**: Symposium selection decision made
- **Feb. 1, 2017**: Call for Papers for accepted Symposia
- **May 15, 2017**: Paper submission due
- **June 30, 2017**: Notification of Acceptance
- **July 22, 2017**: Camera-ready paper due.

SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page



SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

Contents

Volume 33 | Number 6 | November 2016

SPECIAL SECTION

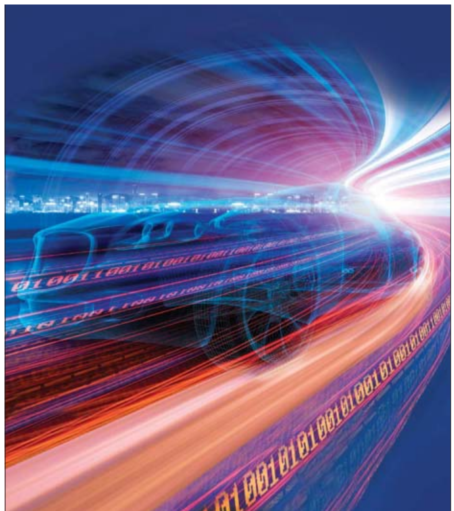
SIGNAL PROCESSING FOR SMART VEHICLE TECHNOLOGIES

12 FROM THE GUEST EDITORS
John H.L. Hansen, Kazuya Takeda, Sanjeev M. Naik, Mohan M. Trivedi, Gerhard U. Schmidt, and Yingying (Jennifer) Chen

14 DRIVER-BEHAVIOR MODELING USING ON-ROAD DRIVING DATA
Chiyomi Miyajima and Kazuya Takeda

22 DRIVER STATUS MONITORING SYSTEMS FOR SMART VEHICLES USING PHYSIOLOGICAL SENSORS
Youjun Choi, Sang Ik Han, Seung-Hyun Kong, and Hyunwoo Ko

PG. 123



ON THE COVER

Signal processing is playing an increasingly substantial role in advancing today's vehicles into "smart" ones. The special section in this issue provides a venue for summarizing, educating, and sharing the state of the art in signal processing applied to automotive systems.

COVER IMAGE: ©ISTOCKPHOTO.COM/NADLA

35 SMART DRIVER MONITORING: WHEN SIGNAL PROCESSING MEETS HUMAN FACTORS
Amirhossein S. Aghaei, Birsen Donmez, Cheng Chen Liu, Dengbo He, George Liu, Konstantinos N. Plataniotis, Huei-Yen Winnie Chen, and Zohreh Sojoudi

49 CONVERSATIONAL IN-VEHICLE DIALOG SYSTEMS
Fuliang Weng, Pongtep Angkititrakul, Elizabeth E. Shriberg, Larry Heck, Stanley Peters, and John H.L. Hansen

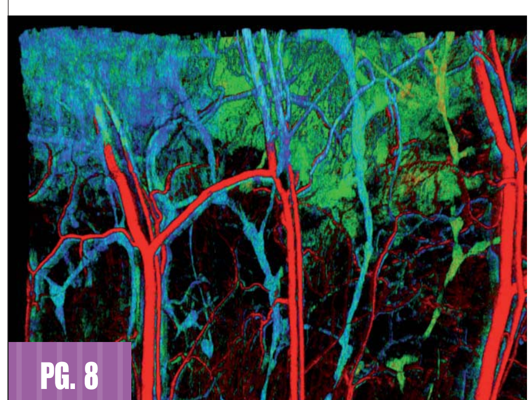
61 RECENT ADVANCES IN ACTIVE NOISE CONTROL INSIDE AUTOMOBILE CABINS
Prasanga N. Samarasinghe, Wen Zhang, and Thushara D. Abhayapala

74 COORDINATION OF COOPERATIVE AUTONOMOUS VEHICLES
Robert Hult, Gabriel R. Campos, Erik Steinmetz, Lars Hammarstrand, Paolo Falcone, and Henk Wymeersch

FEATURES

85 THE RACE TO IMPROVE RADAR IMAGERY
Lifan Zhao, Lu Wang, Lei Yang, Abdelhak M. Zoubir, and Guoan Bi

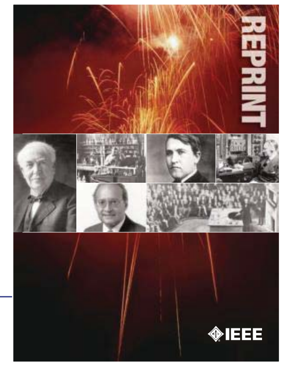
103 MATCHING THEORY
Siavash Bayat, Yonghui Li, Lingyang Song, and Zhu Han



PG. 8

IEEE SIGNAL PROCESSING MAGAZINE (ISSN 1053-5888) (ISPREG) is published bimonthly by the Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, 17th Floor, New York, NY 10016-5997 USA (+1 212 419 7900). Responsibility for the contents rests upon the authors and not the IEEE, the Society, or its members. Annual member subscriptions included in Society fee. Nonmember subscriptions available upon request. **Individual copies:** IEEE Members US$20.00 (first copy only), nonmembers US$213.00 per copy. 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.2016.2610483



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: ______________

IEEE Advancing Technology for Humanity

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
2. Engineering management
3. Project engineering management
4. Research and development management
5. Design engineering management —analog
6. Design engineering management —digital
7. Research and development engineering
8. Design/development engineering —analog
9. Design/development engineering—digital
10. Hardware engineering
11. Software design/development
12. Computer science
13. Science/physics/mathematics
14. Engineering (not elsewhere specified)
15. Marketing/sales/purchasing
16. Consulting
17. Education/teaching
18. Retired
19. Other________________

C. Principal responsibility → ☐

1. Engineering and scientific management
2. Management other than engineering
3. Engineering design
4. Engineering
5. Software: science/mngmnt/engineering
6. Education/teaching
7. Consulting
8. Retired
9. Other________________

D. Title → ☐

1. Chairman of the Board/President/CEO
2. Owner/Partner
3. General Manager
4. VP Operations
5. VP Engineering/Dir. Engineering
6. Chief Engineer/Chief Scientist
7. Engineering Management
8. Scientific Management
9. Member of Technical Staff
10. Design Engineering Manager
11. Design Engineer
12. Hardware Engineer
13. Software Engineer
14. Computer Scientist
15. Dean/Professor/Instructor
16. Consultant
17. Retired
18. Other________________

Are you now or were you ever a member of IEEE?
☐ Yes ☐ No If yes, provide, if known:

Membership Number Grade Year Expired

4 Please Sign Your Application

I hereby apply for IEEE membership and agree to be governed by the IEEE Constitution, Bylaws, and Code of Ethics. I understand that IEEE will communicate with me regarding my individual membership and all related benefits. **Application must be signed.**

Signature Date

Over Please

(continued on next page)

5 Add IEEE Society Memberships (Optional)

The 39 IEEE Societies support your technical and professional interests. Many society memberships include a personal subscription to the core journal, magazine, or newsletter of that society. **For a complete list of everything included with your IEEE Society membership, visit www.ieee.org/join.** All prices are quoted in US dollars.

Please check ✔ the appropriate box.

Society	Code	BETWEEN 16 AUG 2016-28 FEB 2017 PAY	BETWEEN 1 MAR 2017-15 AUG 2017 PAY
IEEE Aerospace and Electronic Systems ⌘ ⍓	AES010	25.00 ☐	12.50 ☐
IEEE Antennas and Propagation ⌘ ⍓	AP003	15.00 ☐	7.50 ☐
IEEE Broadcast Technology ▪ ⌘ ⍓	BT002	15.00 ☐	7.50 ☐
IEEE Circuits and Systems ▪ ⌘ ⍓	CAS004	22.00 ☐	11.00 ☐
IEEE Communications ⌘ ⍓	COM019	33.00 ☐	16.50 ☐
IEEE Components, Packaging, & Manu. Tech. ▪ ⌘ ⍓	CPMT021	15.00 ☐	7.50 ☐
IEEE Computational Intelligence ⌘ ⍓	CIS011	29.00 ☐	14.50 ☐
IEEE Computer ▪ ⌘ ⍓	C016	60.00 ☐	30.00 ☐
IEEE Consumer Electronics ▪ ⌘ ⍓	CE008	20.00 ☐	10.00 ☐
IEEE Control Systems ⌘ ⍓	CS023	25.00 ☐	12.50 ☐
IEEE Dielectrics and Electrical Insulation ⌘ ⍓	DEI032	26.00 ☐	13.00 ☐
IEEE Education ⌘ ⍓	E025	20.00 ☐	10.00 ☐
IEEE Electromagnetic Compatibility ▪ ⌘ ⍓	EMC027	31.00 ☐	15.50 ☐
IEEE Electron Devices ▪ ⌘ ⍓	ED015	18.00 ☐	9.00 ☐
IEEE Engineering in Medicine and Biology ⌘ ⍓	EMB018	40.00 ☐	20.00 ☐
IEEE Geoscience and Remote Sensing ⌘ ⍓	GRS029	19.00 ☐	9.50 ☐
IEEE Industrial Electronics ⌘ ⍓	IE013	9.00 ☐	4.50 ☐
IEEE Industry Applications ⌘ ⍓	IA034	20.00 ☐	10.00 ☐
IEEE Information Theory ▪ ⌘ ⍓	IT012	30.00 ☐	15.00 ☐
IEEE Instrumentation and Measurement ⌘ ⍓	IM009	29.00 ☐	14.50 ☐
IEEE Intelligent Transportation Systems ▪ ⌘ ⍓	ITSS038	35.00 ☐	17.50 ☐
IEEE Magnetics ▪ ⍓ ♦⌘	MAG033	26.00 ☐	13.00 ☐
IEEE Microwave Theory and Techniques ⌘ ⍓	MTT017	24.00 ☐	12.00 ☐
IEEE Nuclear and Plasma Sciences ▪ ⌘ ⍓	NPS005	35.00 ☐	17.50 ☐
IEEE Oceanic Engineering ▪ ⌘ ⍓	OE022	19.00 ☐	9.50 ☐
IEEE Photonics ▪ ⌘ ⍓	PHO036	34.00 ☐	17.00 ☐
IEEE Power Electronics ⌘ ⍓	PEL035	25.00 ☐	12.50 ☐
IEEE Power & Energy ▪ ⌘ ⍓	PE031	35.00 ☐	17.50 ☐
IEEE Product Safety Engineering ▪	PSE043	35.00 ☐	17.50 ☐
IEEE Professional Communication ▪ ⌘ ⍓	PC026	31.00 ☐	15.50 ☐
IEEE Reliability ▪ ⌘ ⍓	RL007	35.00 ☐	17.50 ☐
IEEE Robotics and Automation ⌘ ⍓	RA024	9.00 ☐	4.50 ☐
IEEE Signal Processing ▪ ⌘ ⍓	SP001	22.00 ☐	11.00 ☐
IEEE Social Implications of Technology ▪ ⌘ ⍓	SIT030	33.00 ☐	16.50 ☐
IEEE Solid-State Circuits ⌘ ⍓	SSC037	22.00 ☐	11.00 ☐
IEEE Systems, Man, & Cybernetics ⌘ ⍓	SMC028	12.00 ☐	6.00 ☐
IEEE Technology & Engineering Management ⌘ ⍓	TEM014	35.00 ☐	17.50 ☐
IEEE Ultrasonics, Ferroelectrics, & Frequency Control ▪ ⌘ ⍓	UFFC020	20.00 ☐	10.00 ☐
IEEE Vehicular Technology ▪ ⌘ ⍓	VT006	18.00 ☐	9.00 ☐

Legend—Society membership includes:

- ⌘ One or more Society publications
- ▪ Society newsletter
- ⍓ Online access to publication
- ♦ CD-ROM of selected society publications

Complete both sides of this form, sign, and return to:
IEEE MEMBERSHIP APPLICATION PROCESSING
445 HOES LN, PISCATAWAY, NJ 08854-4141 USA
or fax to +1 732 981 0225
or join online at www.ieee.org/join

Please reprint your full name here ______________________

6 2017 IEEE Membership Rates (student rates available online)

IEEE member dues and regional assessments are based on where you live and when you apply. Membership is based on the calendar year from 1 January through 31 December. All prices are quoted in US dollars.

Please check ✔ the appropriate box.

RESIDENCE	BETWEEN 16 AUG 2016-28 FEB 2017 PAY	BETWEEN 1 MAR 2017-15 AUG 2017 PAY
United States	$199.00 ☐	$99.50 ☐
Canada (NB, NF, NS, and PEI HST)*	$190.35 ☐	$95.18 ☐
Canada (ON HST)	$187.37 ☐	$93.69 ☐
Canada (GST)*	$175.45 ☐	$87.73 ☐
Canada (GST and QST Quebec)	$190.31 ☐	$95.16 ☐
Africa, Europe, Middle East	$162.00 ☐	$81.00 ☐
Latin America	$153.00 ☐	$76.50 ☐
Asia, Pacific	$154.00 ☐	$77.00 ☐

*IEEE Canada Business No. 125634188

Minimum Income or Unemployed Provision

Applicants who certify that their prior year income did not exceed US$14,900 (or equivalent) or were not employed are granted 50% reduction in: full-year dues, regional assessment and fees for one IEEE Membership plus one Society Membership. If applicable, please check appropriate box and adjust payment accordingly. Student members are not eligible.

☐ I certify I earned less than US$14,900 in 2016
☐ I certify that I was unemployed in 2016

7 More Recommended Options

Proceedings of the IEEE .... print $49.00 ☐ **or** online $43.00 ☐
Proceedings of the IEEE (print/online combination) .... $59.00 ☐
IEEE Standards Association (IEEE-SA) .... $54.00 ☐
IEEE Women in Engineering (WIE) .... $25.00 ☐

8 Payment Amount

Please total the Membership dues, Society dues, and other amounts from this page:

IEEE Membership dues ⑥ .... $______
IEEE Society dues (optional) ⑤ .... $______
IEEE-SA/WIE dues (optional) ⑦ .... $______
Proceedings of the IEEE (optional) ⑦ .... $______
Canadian residents pay 5% GST or appropriate HST (BC–12%; ON–13%; NB, NF, NS, PEI–15%) on Society payments & publications only .... TAX $______
AMOUNT PAID .... **TOTAL** $______

Payment Method

All prices are quoted in US dollars. You may pay for IEEE membership by credit card (see below), check, or money order payable to IEEE, drawn on a US bank.

☐ Check ☐ VISA ☐ MasterCard ☐ Diners Club International ☐ American Express ☐ Discover

Credit Card Number

MONTH YEAR
EXPIRATION DATE

CARDHOLDER'S 5-DIGIT ZIP CODE
(BILLING STATEMENT ADDRESS) USA ONLY

Name as it appears on card

Signature

Auto Renew my Memberships and Subscriptions (available when paying by credit card).
☐ I agree to the Terms and Conditions located at www.ieee.org/autorenew

9 Were You Referred to IEEE?

☐ Yes ☐ No If yes, provide the following:
Member Recruiter Name______________________
IEEE Recruiter's Member Number (Required)______________________

CAMPAIGN CODE | PROMO CODE

16-MEM-031 P 6/16

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

Information for Authors

(Updated/Effective January 2015)

For Transactions and Journals:

Authors are encouraged to submit manuscripts of Regular papers (papers which provide a complete disclosure of a technical premise), or Comment Correspondences (brief items that provide comment on a paper previously published in these TRANSACTIONS).

Submissions/resubmissions must be previously unpublished and may not be under consideration elsewhere.

Every manuscript must:

i. provide a clear statement of the problem and what the contribution of the work is to the relevant research community;
ii. state why this contribution is significant (what impact it will have);
iii. provide citation of the published literature most closely related to the manuscript; and
iv. state what is distinctive and new about the current manuscript relative to these previously published works.

By submission of your manuscript to these TRANSACTIONS, all listed authors have agreed to the authorship list and all the contents and confirm that the work is original and that figures, tables and other reported results accurately reflect the experimental work. In addition, the authors all acknowledge that they accept the rules established for publication of manuscripts, including agreement to pay all overlength page charges, color charges, and any other charges and fees associated with publication of the manuscript. Such charges are not negotiable and cannot be suspended. The corresponding author is responsible for obtaining consent from all co-authors and, if needed, from sponsors before submission.

In order to be considered for review, a paper must be within the scope of the journal and represent a novel contribution. A paper is a candidate for an Immediate Rejection if it is of limited novelty, e.g. a straightforward combination of theories and algorithms that are well established and are repeated on a known scenario. Experimental contributions will be rejected without review if there is insufficient experimental data. These TRANSACTIONS are published in English. Papers that have a large number of typographical and/or grammatical errors will also be rejected without review.

In addition to presenting a novel contribution, acceptable manuscripts must describe and cite related work in the field to put the contribution in context. Do not give theoretical derivations or algorithm descriptions that are easily found in the literature; merely cite the reference.

New and revised manuscripts should be prepared following the "Manuscript Submission" guidelines below, and submitted to the online manuscript system, ScholarOne Manuscripts. Do not send original submissions or revisions directly to the Editor-in-Chief or Associate Editors; they will access your manuscript electronically via the ScholarOne Manuscript system.

Manuscript Submission. Please follow the next steps.

1. *Account in ScholarOne Manuscripts.* If necessary, create an account in the on-line submission system ScholarOne Manuscripts. Please check first if you already have an existing account which is based on your e-mail address and may have been created for you when you reviewed or authored a previous paper.
2. *Electronic Manuscript.* Prepare a PDF file containing your manuscript in double-column, single-spaced format using a font size of 10 points or larger, having a margin of at least 1 inch on all sides. Upload this version of the manuscript as a PDF file "double.pdf" to the ScholarOne-Manuscripts site. Since many reviewers prefer a larger font, you are strongly encouraged to also submit a single-column, double-spaced version (11 point font or larger), which is easy to create with the templates provided **IEEE Author Digital Toolbox (http://www.ieee.org/publications_standards/publications/authors/authors_journals.html)**. Page length restrictions will be determined by the double-column version. Proofread your submission, confirming that all figures and equations are visible in your document before you "SUBMIT" your manuscript. Proofreading is critical; once you submit your manuscript, the manuscript cannot be changed in any way. You may also submit your manuscript as a .PDF or MS Word file. The system has the capability of converting your files to PDF, however it is your responsibility to confirm that the conversion is correct and there are no font or graphics issues prior to completing the submission process.
3. *EDICS (Not applicable to Journal of Selected Topics in Signal Processing).* All submissions must be classified by the author with an EDICS (Editors' Information Classification Scheme) selected from the list of EDICS published online at the at the publication's EDICS webpage (*please see the list below). Upon submission of a new manuscript, please choose the EDICS categories that best suit your manuscript. Failure to do so will likely result in a delay of the peer review process.
4. *Additional Documents for Review.* Please upload pdf versions of all items in the reference list that are not publicly available, such as unpublished (submitted) papers. Graphical abstracts and supplemental materials intended to appear with the final paper (see below) must also be uploaded for review at the time of the initial submission for consideration in the review process. Use short filenames without spaces or special characters. When the upload of each file is completed, you will be asked to provide a description of that file.
5. *Supplemental Materials.* IEEE Xplore can publish multimedia files (audio, images, video), datasets, and software (e.g. Matlab code) along with your paper. Alternatively, you can provide the links to such files in a README file that appears on Xplore along with your paper. For details, please see IEEE Author Digital Toolbox under "Multimedia." To make your work reproducible by others, these TRANSACTIONS encourages you to submit all files that can recreate the figures in your paper.
6. *Submission.* After uploading all files and proofreading them, submit your manuscript by clicking "Submit." A confirmation of the successful submission will open on screen containing the manuscript tracking number and will be followed with an e-mail confirmation to the corresponding and all contributing authors. Once you click "Submit," your manuscript cannot be changed in any way.
7. *Copyright Form and Consent Form.* By policy, IEEE owns the copyright to the technical contributions it publishes on behalf of the interests of the IEEE, its authors, and their employers; and to facilitate the appropriate reuse of this material by others. To comply with the IEEE copyright policies, authors are required to sign and submit a completed "IEEE Copyright and Consent Form" prior to publication by the IEEE. The IEEE recommends authors to use an effective electronic copyright form (eCF) tool within the ScholarOne Manuscripts system. You will be redirected to the "IEEE Electronic Copyright Form" wizard at the end of your original submission; please simply sign the eCF by typing your name at the proper location and click on the "Submit" button.

Comment Correspondence. Comment Correspondences provide brief comments on material previously published in these TRANSACTIONS. These items may not exceed 2 pages in double-column, single spaced format, using 9 point type, with margins of 1 inch minimum on all sides, and including: title, names and contact information for authors, abstract, text, references, and an appropriate number of illustrations and/or tables. Correspondence items are submitted in the same way as regular manuscripts (see "Manuscript Submission" above for instructions).

Authors may also submit manuscripts of overview articles, but note that these include an additional white paper approval process http://www.signalprocessingsociety.org/publications/overview-articles/. *[This does not apply to the Journal of Selected Topics in Signal Processing. Please contact the Editor-in-Chief.]*

Digital Object Identifier

Manuscript Length. For the initial submission of a regular paper, the manuscript may not exceed 13 double-column pages (10 point font), including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs; and all references. Supplemental materials and graphical abstracts are not included in the page count. For regular papers, the revised manuscript may not exceed 16 double-column pages (10 point font), including title; names of authors and their complete contact information; abstract; text; all images, figures and tables, appendices and proofs; and all references. For Overview Papers, the maximum length is double that for regular submissions at each stage (please reference http://www.signalprocessingsociety.org/publications/overview-articles/ for more information).

Note that any paper in excess of 10 pages will be subject to mandatory overlength page charges. Since changes recommended as a result of peer review may require additions to the manuscript, it is strongly recommended that you practice economy in preparing original submissions. Note: Papers submitted to the TRANSACTIONS ON MULTIMEDIA in excess of 8 pages will be subject to mandatory overlength page charges.

Exceptions to manuscript length requirements may, under extraordinary circumstances, be granted by the Editor-in-Chief. However, such exception does not obviate your requirement to pay any and all overlength or additional charges that attach to the manuscript.

Resubmission of Previously Rejected Manuscripts. Authors of manuscripts rejected from any journal are allowed to resubmit their manuscripts only once. At the time of submission, you will be asked whether your manuscript is a new submission or a resubmission of an earlier rejected manuscript. If it is a resubmission of a manuscript previously rejected by any journal, you are expected to submit supporting documents identifying the previous submission and detailing how your new version addresses all of the reviewers' comments. Papers that do not disclose connection to a previously rejected paper or that do not provide documentation as to changes made may be immediately rejected.

Author Misconduct. Author misconduct includes plagiarism, self-plagiarism, and research misconduct, including falsification or misrepresentation of results. All forms of misconduct are unacceptable and may result in sanctions and/or other corrective actions. Plagiarism includes copying someone else's work without appropriate credit, using someone else's work without clear delineation of citation, and the uncited reuse of an author's previously published work that also involves other authors. Self-plagiarism involves the verbatim copying or reuse of an authors own prior work without appropriate citation, including duplicate submission of a single journal manuscript to two different journals, and submission of two different journal manuscripts which overlap substantially in language or technical contribution. For more information on the definitions, investigation process, and corrective actions related to author misconduct, see the Signal Processing Society Policies and Procedures Manual, Section 6.1. http://www.signalprocessingsociety.org/about-sps/governance/policy-procedure/part-2. Author misconduct may also be actionable by the IEEE under the rules of Member Conduct.

Extensions of the Author's Prior Work. It is acceptable for conference papers to be used as the basis for a more fully developed journal submission. Still, authors are required to cite their related prior work; the papers cannot be identical; and the journal publication must include substantively novel aspects such as new experimental results and analysis or added theoretical work. The journal publication should clearly specify how the journal paper offers novel contributions when citing the prior work. Limited overlap with prior journal publications with a common author is allowed only if it is necessary for the readability of the paper, and the prior work must be cited as the primary source.

Submission Format. Authors are required to prepare manuscripts employing the on-line style files developed by IEEE, which include guidelines for abbreviations, mathematics, and graphics. All manuscripts accepted for publication will require the authors to make final submission employing these style files. The style files are available on the web at the **IEEE Author Digital Toolbox** under "Template for all TRANSACTIONS." (LaTeX and MS Word). Please note the following requirements about the abstract:

- The abstract must be a concise yet comprehensive reflection of what is in your article.
- The abstract must be self-contained, without abbreviations, footnotes, displayed equations, or references.
- The abstract must be between 150-250 words.
- The abstract should include a few keywords or phrases, as this will help readers to find it. Avoid over-repetition of such phrases as this can result in a page being rejected by search engines.

In addition to written abstracts, papers may include a graphical abstract; see http://www.ieee.org/publications_standards/publications/authors/authors_journals.html. for options and format requirements.

IEEE supports the publication of author names in the native language alongside the English versions of the names in the author list of an article. For more information, see "Author names in native languages" (http://www.ieee.org/publications_standards/publications/authors/auth_names_native_lang.pdf) on the IEEE Author Digital Toolbox page.

Open Access. The publication is a hybrid journal, allowing either Traditional manuscript submission or Open Access (author-pays OA) manuscript submission. Upon submission, if you choose to have your manuscript be an Open Access article, you commit to pay the discounted $1,750 OA fee if your manuscript is accepted for publication in order to enable unrestricted public access. Any other application charges (such as overlength page charge and/or charge for the use of color in the print format) will be billed separately once the manuscript formatting is complete but prior to the publication. If you would like your manuscript to be a Traditional submission, your article will be available to qualified subscribers and purchasers via IEEE Xplore. No OA payment is required for Traditional submission.

Page Charges.

Voluntary Page Charges. Upon acceptance of a manuscript for publication, the author(s) or his/her/their company or institution will be asked to pay a charge of $110 per page to cover part of the cost of publication of the first ten pages that comprise the standard length (two pages, in the case of Correspondences).

Mandatory Page Charges The author(s) or his/her/their company or institution will be billed $220 per each page in excess of the first ten published pages for regular papers and six published pages for correspondence items. (**NOTE: Papers accepted to IEEE TRANSACTIONS ON MULTIMEDIA in excess of 8 pages will be subject to mandatory overlength page charges.) These are mandatory page charges and the author(s) will be held responsible for them. They are not negotiable or voluntary. The author(s) signifies his willingness to pay these charges simply by submitting his/her/their manuscript to the TRANSACTIONS. The Publisher holds the right to withhold publication under any circumstance, as well as publication of the current or future submissions of authors who have outstanding mandatory page charge debt. No mandatory overlength page charges will be applied to overview articles in the Society's journals.

Color Charges. Color figures which appear in color only in the electronic (Xplore) version can be used free of charge. In this case, the figure will be printed in the hardcopy version in grayscale, and the author is responsible that the corresponding grayscale figure is intelligible. Color reproduction charges for print are the responsibility of the author. Details of the associated charges can be found on the IEEE Publications page.

Payment of fees on color reproduction is not negotiable or voluntary, and the author's agreement to publish the manuscript in these TRANSACTIONS is considered acceptance of this requirement.

*EDICS Webpages:

IEEE TRANSACTIONS ON SIGNAL PROCESSING:
http://www.signalprocessingsociety.org/publications/periodicals/tsp/TSP-EDICS/
IEEE TRANSACTIONS ON IMAGE PROCESSING:
http://www.signalprocessingsociety.org/publications/periodicals/image-processing/tip-edics/
IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE / ACM:
http://www.signalprocessingsociety.org/publications/periodicals/taslp/taslp-edics/
IEEE TRANSACTIONS ON INFORMATION, FORENSICS AND SECURITY:
http://www.signalprocessingsociety.org/publications/periodicals/forensics/forensics-edics/
IEEE TRANSACTIONS ON MULTIMEDIA:
http://www.signalprocessingsociety.org/tmm/tmm-edics/
IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING:
http://www.signalprocessingsociety.org/publications/periodicals/tci/tci-edics/
IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS:
http://www.signalprocessingsociety.org/publications/periodicals/tsipn/tsipn-edics/

IEEE SignalProcessing Previous Page | Contents | Zoom in | Zoom out | Front Cover | Search Issue | Next Page Qmags THE WORLD'S NEWSSTAND®

2016 IEEE SIGNAL PROCESSING SOCIETY MEMBERSHIP APPLICATION

Mail to: IEEE OPERATIONS CENTER, ATTN: Matthew Plotner, Member and Geographic Activities, 445 Hoes Lane, Piscataway, New Jersey 08854 USA
or Fax to (732) 981-0225 (credit card payments only.)
For info call (732) 981-0060 or 1 (800) 678-IEEE or E-mail: new.membership@ieee.org

IEEE

1. PERSONAL INFORMATION

NAME AS IT SHOULD APPEAR ON IEEE MAILINGS: **SEND MAIL TO**: ☐ **Home Address** OR ☐ **Business/School Address**
If not indicated, mail will be sent to home address. Note: Enter your name as you wish it to appear on membership card and all correspondence.
PLEASE PRINT Do not exceed 40 characters or spaces per line. Abbreviate as needed. Please circle your last/surname as a key identifier for the IEEE database.

TITLE ______ FIRST OR GIVEN NAME ______ MIDDLE NAME ______ SURNAME/LAST NAME

HOME ADDRESS

CITY ______ STATE/PROVINCE ______ POSTAL CODE ______ COUNTRY

2.

Are you now or were you ever a member of IEEE? ☐ **Yes** ☐ **No**
If yes, please provide, if known:

MEMBERSHIP NUMBER ______

Grade ______ **Year Membership Expired:** ______

3. BUSINESS/PROFESSIONAL INFORMATION

Company Name

Department/Division

Title/Position ______ Years in Current Position

Years in the Profession Since Graduation ______ ☐ PE State/Province

Street Address

City ______ State/Province ______ Postal Code ______ Country

4. EDUCATION

A baccalaureate degree from an IEEE recognized educational program assures assignment of “Member” grade. For others, additional information and references may be necessary for grade assignment.

A.
Baccalaureate Degree Received ______ Program/Course of Study

College/University ______ Campus

State/Province ______ Country ______ Mo./Yr. Degree Received

B.
Highest Technical Degree Received ______ Program/Course of Study

College/University ______ Campus

State/Province ______ Country ______ Mo./Yr. Degree Received

5.

Full signature of applicant

6. DEMOGRAPHIC INFORMATION – ALL APPLICANTS -

Date Of Birth ______ ☐ Male ☐ Female
Day Month Year

7. CONTACT INFORMATION

Office Phone/Office Fax ______ Home Phone/Home Fax

Office E-Mail ______ Home E-Mail

8. 2016 IEEE MEMBER RATES

IEEE DUES Residence	16 Aug 14-28 Feb 15 Pay Full Year	1 Mar -15 Aug 15 Pay Half Year**
United States	$197.00 ☐	$98.50 ☐
Canada (incl. GST)	$173.35 ☐	$86.68 ☐
Canada (incl. HST for PEI)	$186.58 ☐	$93.29 ☐
Canada (incl. HST for Nova Scotia)	$188.05 ☐	$94.03 ☐
Canada (incl. HST for NB, NF and ON)	$185.11 ☐	$92.56 ☐
Canada (incl. GST and QST Quebec)	$188.01 ☐	$94.01 ☐
Africa, Europe, Middle East	$160.00 ☐	$80.00 ☐
Latin America	$151.00 ☐	$75.50 ☐
Asia, Pacific	$152.00 ☐	$76.00 ☐

Canadian Taxes (GST/HST): All supplies, which include dues, Society membership fees, online products and publications (except CD-ROM and DVD media), shipped to locations within Canada are subject to the GST of 5% or the HST of 12%, 13% or 15%, depending on the Province to which the materials are shipped. GST and HST do not apply to Regional Assessments. (IEEE Canadian Business Number 12563 4188 RT0001)
Value Added Tax (VAT) in the European Union: In accordance with the European Union Council Directives 2002/38/EC and 77/388/EEC amended by Council Regulation (EC)792/2002, IEEE is required to charge and collect VAT on electronic/digitized products sold to private consumers that reside in the European Union. The VAT rate applied is the EU member country standard rate where the consumer is resident. (IEEE's VAT registration number is EU826000081)
U.S. Sales Taxes: Please add applicable state and local sales and use tax on orders shipped to **Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Illinois, Indiana, Kentucky, Massachusetts, Maryland, Michigan, Minnesota, Missouri, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, West Virginia, Wisconsin.** Customers claiming a tax exemption must include an appropriate and properly completed tax-exemption certificate with their first order.

2016 SPS MEMBER RATES

		16 Aug-28 Feb Pay Full Year	1 Mar-15 Aug Pay Half Year
Signal Processing Society Membership Fee*		$ 22.00 ☐	$ 11.00 ☐

Fee includes: **IEEE Signal Processing Magazine** (electronic and digital), **Inside Signal Proc. eNewsletter** (electronic) and **IEEE Signal Processing Society Content Gazette** (electronic).

Add $17 to enhance SPS Membership and also receive:		$17.00 ☐	$ 8.50 ☐

IEEE Signal Processing Magazine (print) and **SPS Digital Library**: online access to Signal Processing Magazine, Signal Processing Letters, Journal of Selected Topics in Signal Processing, Trans. on Audio, Speech, and Language Processing, Trans. on Image Processing, Trans. on Information Forensics and Security and Trans. on Signal Processing.

Publications available only with SPS membership:

Publication	Format	Pay Full Year	Pay Half Year
Signal Processing, IEEE Transactions on:	Print	$209.00 ☐	$104.50 ☐
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print	$160.00 ☐	$ 80.00 ☐
Image Processing, IEEE Transactions on:	Print	$207.00 ☐	$103.50 ☐
Information Forensics and Security, IEEE Trans. on:	Print	$179.00 ☐	$ 89.50 ☐
IEEE Journal of Selected Topics in Signal Processing:	Print	$176.00 ☐	$ 88.00 ☐
Affective Computing, IEEE Transactions on:	Electronic	$ 36.00 ☐	$ 18.00 ☐
Biomedical and Health Informatics, IEEE Journal of:	Print	$ 55.00 ☐	$ 27.50 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$ 65.00 ☐	$ 32.50 ☐
IEEE Cloud Computing	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
IEEE Trans. on Cognitive Comm. &Networking	Electronic	$ 27.00 ☐	$ 13.50 ☐
IEEE Trans. on Computational Imaging	Electronic	$ 28.00 ☐	$ 14.00 ☐
IEEE Trans. on Big Data	Electronic	$ 26.00 ☐	$ 13.00 ☐
IEEE Trans. on Molecular, Biological, & Multi-scale Communications	Electronic	$ 25.00 ☐	$ 12.50 ☐
IEEE Internet of Things Journal	Electronic	$ 26.00 ☐	$ 13.00 ☐
IEEE Trans. on Cloud Computing	Electronic	$ 43.00 ☐	$ 21.50 ☐
IEEE Trans. on Computational Social Systems	Electronic	$ 30.00 ☐	$ 15.00 ☐
IEEE Trans. on Signal & Info Proc. Over Networks	Electronic	$ 28.00 ☐	$ 14.00 ☐
IEEE Biometrics Compendium:	Online	$ 30.00 ☐	$ 15.00 ☐
Computing in Science & Engrg. Mag.:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$ 69.00 ☐	$ 34.50 ☐
Medical Imaging, IEEE Transactions on:	Print	$ 74.00 ☐	$ 37.00 ☐
	Electronic	$ 53.00 ☐	$ 26.50 ☐
	Print & Electronic	$ 89.00 ☐	$ 44.50 ☐
Mobile Computing, IEEE Transactions on:	ELE/Print Abstract/CD-ROM	$ 41.00 ☐	$ 20.50 ☐
Multimedia, IEEE Transactions on:	Electronic	$ 43.00 ☐	$ 21.50 ☐
IEEE MultiMedia Magazine:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$ 69.00 ☐	$ 34.50 ☐
Network Science and Engrg., IEEE Trans. on:	Electronic	$ 34.00 ☐	$ 17.00 ☐
IEEE Reviews in Biomedical Engineering:	Print	$ 25.00 ☐	$ 12.50 ☐
	Print & Electronic	$ 40.00 ☐	$ 20.00 ☐
IEEE Security and Privacy Magazine:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$ 69.00 ☐	$ 34.50 ☐
IEEE Sensors Journal:	Electronic	$ 40.00 ☐	$ 20.00 ☐
Smart Grid, IEEE Transactions on:	Print	$100.00 ☐	$ 50.00 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$120.00 ☐	$ 60.00 ☐
Wireless Communications, IEEE Transactions on:	Print	$127.00 ☐	$ 63.50 ☐
	Electronic	$ 49.00 ☐	$ 24.50 ☐
	Print & Electronic	$127.00 ☐	$ 63.50 ☐
IEEE Wireless Communications Letters:	Electronic	$ 19.00 ☐	$ 9.50 ☐

*IEEE membership required or requested
Affiliate application to join SP Society only. Amount Paid $______

9.

IEEE Membership Affiliate Fee (See pricing in Section 8) $______

Signal Processing Society Fees $______

Canadian residents pay 5% GST or 13% HST
Reg. No. 125634188 on Society payment(s) & pubs only Tax $______

AMOUNT PAID WITH APPLICATION TOTAL $______
Prices subject to change without notice.
☐ **Check or money order enclosed Payable to IEEE on a U.S. Bank**
☐ **American Express** ☐ **VISA** ☐ **MasterCard**
☐ **Diners Club**

Exp. Date/ Mo./Yr.

Cardholder Zip Code Billing Statement Address/USA Only

Full signature of applicant using credit card ______ Date

10. WERE YOU REFERRED?

☐ Yes ☐ No If yes, please provide the follow information:
Member Recruiter Name: ______
IEEE Recruiter's Member Number (Required): ______

IEEE Signal Processing Society

2016 IEEE SIGNAL PROCESSING SOCIETY STUDENT MEMBERSHIP APPLICATION

(Current and reinstating IEEE members joining SPS complete areas 1, 2, 8, 9.)

Mail to: **IEEE OPERATIONS CENTER, ATTN: Matthew Plotner, Member and Geographic Activities, 445 Hoes Lane, Piscataway, New Jersey 08854 USA**
or Fax to (732) 981-0225 (credit card payments only.)
For info call (732) 981-0060 or 1 (800) 678-IEEE or E-mail: new.membership@ieee.org

IEEE

1. PERSONAL INFORMATION

NAME AS IT SHOULD APPEAR ON IEEE MAILINGS: SEND MAIL TO: ☐ **Home Address** OR ☐ **Business/School Address**
If not indicated, mail will be sent to home address. Note: Enter your name as you wish it to appear on membership card and all correspondence.
PLEASE PRINT Do not exceed 40 characters or spaces per line. Abbreviate as needed. Please circle your last/surname as a key identifier for the IEEE database.

TITLE | FIRST OR GIVEN NAME | MIDDLE NAME | SURNAME/LAST NAME

HOME ADDRESS

CITY | STATE/PROVINCE | POSTAL CODE | COUNTRY

2.

Are you now or were you ever a member of IEEE? ☐ **Yes** ☐ **No**
If yes, please provide, if known:

MEMBERSHIP NUMBER______________________

Grade____________**Year Membership Expired:**______________

3. BUSINESS/PROFESSIONAL INFORMATION

Company Name

Department/Division

Title/Position | Years in Current Position

Years in the Profession Since Graduation | ☐ PE State/Province

Street Address

City | **State/Province** | **Postal Code** | **Country**

4. EDUCATION

A baccalaureate degree from an IEEE recognized educational program assures assignment of "Member" grade. For others, additional information and references may be necessary for grade assignment.

A.

Baccalaureate Degree Received | Program/Course of Study

College/University | Campus

State/Province | Country | Mo./Yr. Degree Received

B.

Highest Technical Degree Received | Program/Course of Study

College/University | Campus

State/Province | Country | Mo./Yr. Degree Received

5.

Full signature of applicant

6. DEMOGRAPHIC INFORMATION – ALL APPLICANTS -

Date Of Birth ____________________ ☐ Male ☐ Female
Day Month Year

7. CONTACT INFORMATION

Office Phone/Office Fax | Home Phone/Home Fax

Office E-Mail | Home E-Mail

8. 2016 IEEE STUDENT MEMBER RATES

IEEE DUES Residence	16 Aug 14-28 Feb 15 Pay Full Year	1 Mar -15 Aug 15 Pay Half Year**
United States	$32.00 ☐	$16.00 ☐
Canada (incl. GST)	$33.60 ☐	$16.80 ☐
Canada (incl. HST for NB, NF, and ON)	$36.16 ☐	$18.08 ☐
Canada (incl. HST for Nova Scotia)	$36.80 ☐	$18.40 ☐
Canada (incl. HST for PEI)	$36.48 ☐	$18.24 ☐
Canada (incl. GST and QST Quebec)	$36.79 ☐	$18.40 ☐
Africa, Europe, Middle East, Latin America, Asia, Pacific	$27.00 ☐	$13.50 ☐

Canadian Taxes (GST/HST): All supplies, which include dues, Society membership fees, online products and publications (except CD-ROM and DVD media), shipped to locations within Canada are subject to the GST of 5% or the HST of 12%,13% or 15%, depending on the Province to which the materials are shipped. GST and HST do not apply to Regional Assessments. (IEEE Canadian Business Number 12563 4188 RT0001)
Value Added Tax (VAT) in the European Union: In accordance with the European Union Council Directives 2002/38/EC and 77/388/EEC amended by Council Regulation (EC)792/2002, IEEE is required to charge and collect VAT on electronic/digitized products sold to private consumers that reside in the European Union. The VAT rate applied is the EU member country standard rate where the consumer is resident. (IEEE's VAT registration number is EU826000081)
U.S. Sales Taxes: Please add applicable state and local sales and use tax on orders shipped to **Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Illinois, Indiana, Kentucky, Massachusetts, Maryland, Michigan, Minnesota, Missouri, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, West Virginia, Wisconsin.** Customers claiming a tax exemption must include an appropriate and properly completed tax-exemption certificate with their first order.

2016 SPS STUDENT MEMBER RATES

		16 Aug-28 Feb Pay Full Year	1 Mar-15 Aug Pay Half Year
Signal Processing Society Membership Fee*		$11.00 ☐	$ 5.50 ☐

Fee includes**: IEEE Signal Processing Magazine** (electronic and digital), **Inside Signal Processing eNewsletter** (electronic) and **IEEE Signal Processing Society Content Gazette** (electronic).

Add $8 to enhance SPS Membership and also receive:		$ 9.00 ☐	$ 4.50 ☐

IEEE Signal Processing Society Magazine (print) and **SPS Digital Library**: online access to Signal Processing Magazine, Signal Processing Letters, Journal of Selected Topics in Signal Processing, Trans. on Audio, Speech, and Language Processing, Trans. on Image Processing, Trans. on Information Forensics and Security and Trans. on Signal Processing.

Publications available only with SPS membership:

Publication		Full Year	Half Year
Signal Processing, IEEE Transactions on:	Print	$105.00 ☐	$ 52.50 ☐
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print	$ 80.00 ☐	$ 40.00 ☐
Image Processing, IEEE Transactions on:	Print	$104.00 ☐	$ 52.00 ☐
Information Forensics and Security, IEEE Trans. on:	Print	$ 90.00 ☐	$ 45.00 ☐
IEEE Journal of Selected Topics in Signal Processing:	Print	$ 88.00 ☐	$ 44.00 ☐
Affective Computing, IEEE Transactions on:	Electronic	$ 18.00 ☐	$ 9.00 ☐
Biomedical and Health Informatics, IEEE Journal of:	Print	$ 28.00 ☐	$ 14.00 ☐
	Electronic	$ 20.00 ☐	$ 10.00 ☐
	Print & Electronic	$ 33.00 ☐	$ 16.50 ☐
IEEE Cloud Computing	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
IEEE Trans. on Cognitive Comm. &Networking	Electronic	$ 14.00 ☐	$ 7.00 ☐
IEEE Trans. on Computational Imaging	Electronic	$ 14.00 ☐	$ 7.00 ☐
IEEE Trans. on Big Data	Electronic	$ 13.00 ☐	$ 6.50 ☐
IEEE Trans. on Molecular, Biological, & Multi-Scale Communications	Electronic	$ 13.00 ☐	$ 6.50 ☐
IEEE Internet of Things Journal	Electronic	$ 13.00 ☐	$ 6.50 ☐
IEEE Trans. on Cloud Computing	Electronic	$ 22.00 ☐	$ 11.00 ☐
IEEE Trans. on Computational Social Systems	Electronic	$ 15.00 ☐	$ 7.50 ☐
IEEE Trans. on Signal & Info Proc. Over Networks	Electronic	$ 14.00 ☐	$ 7.00 ☐
IEEE Biometrics Compendium:	Online	$ 15.00 ☐	$ 7.50 ☐
Computing in Science & Engrg. Mag.:	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
	Print	$ 35.00 ☐	$ 17.50 ☐
Medical Imaging, IEEE Transactions on:	Print	$ 37.00 ☐	$ 18.50 ☐
	Electronic	$ 27.00 ☐	$ 13.50 ☐
	Print & Electronic	$ 45.00 ☐	$ 22.50 ☐
Mobile Computing, IEEE Transactions on:	ELE/Print Abstract/CD-ROM	$ 21.00 ☐	$ 10.50 ☐
Multimedia, IEEE Transactions on:	Electronic	$ 22.00 ☐	$ 11.00 ☐
IEEE MultiMedia Magazine:	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
	Print	$ 35.00 ☐	$ 17.50 ☐
Network Science and Engrg., IEEE Trans. on:	Electronic	$ 17.00 ☐	$ 8.50 ☐
IEEE Reviews in Biomedical Engineering:	Print	$ 13.00 ☐	$ 6.50 ☐
	Print & Electronic	$ 20.00 ☐	$ 10.00 ☐
IEEE Security and Privacy Magazine:	Electronic and Digital	$ 20.00 ☐	$ 10.00 ☐
	Print	$ 35.00 ☐	$ 17.50 ☐
IEEE Sensors Journal:	Electronic	$ 20.00 ☐	$ 10.00 ☐
Smart Grid, IEEE Transactions on:	Print	$ 50.00 ☐	$ 25.00 ☐
	Electronic	$ 20.00 ☐	$ 10.00 ☐
	Print & Electronic	$ 60.00 ☐	$ 30.00 ☐
Wireless Communications, IEEE Transactions on:	Print	$ 64.00 ☐	$ 32.00 ☐
	Electronic	$ 25.00 ☐	$ 12.50 ☐
	Print & Electronic	$ 64.00 ☐	$ 32.00 ☐
IEEE Wireless Communications Letters:	Electronic	$ 10.00 ☐	$ 5.00 ☐

*IEEE membership required or requested
Affiliate application to join SP Society only. Amount Paid **$**_______________

9.

IEEE Membership Fee (See pricing in Section 8) **$**_______________
Signal Processing Society Fees **$**_______________
Canadian residents pay 5% GST or 13% HST
Reg. No. 125634188 on Society payment(s) & pubs only Tax **$**_______________

AMOUNT PAID WITH APPLICATION TOTAL **$**_______________
Prices subject to change without notice.
☐ **Check or money order enclosed Payable to IEEE on a U.S. Bank**
☐ **American Express** ☐ **VISA** ☐ **MasterCard** ☐ **Diners Club**

Exp. Date/ Mo./Yr.

Cardholder Zip Code Billing Statement Address/USA Only

Full signature of applicant using credit card Date

10. WERE YOU REFERRED?

☐ Yes ☐ No If yes, please provide the follow information:
Member Recruiter Name:______________________________
IEEE Recruiter's Member Number (Required): ______________________

2016 IEEE SIGNAL PROCESSING SOCIETY AFFILIATE MEMBERSHIP APPLICATION

Mail to: IEEE OPERATIONS CENTER, ATTN: Matthew Plotner, Member and Geographic Activities, 445 Hoes Lane, Piscataway, New Jersey 08854 USA
or Fax to (732) 981-0225 (credit card payments only.)
For info call (732) 981-0060 or 1 (800) 678-IEEE or E-mail: new.membership@ieee.org

IEEE

1. PERSONAL INFORMATION

NAME AS IT SHOULD APPEAR ON IEEE MAILINGS: **SEND MAIL TO**: ☐ **Home Address** OR ☐ **Business/School Address**
If not indicated, mail will be sent to home address. Note: Enter your name as you wish it to appear on membership card and all correspondence.
PLEASE PRINT Do not exceed 40 characters or spaces per line. Abbreviate as needed. Please circle your last/surname as a key identifier for the IEEE database.

TITLE FIRST OR GIVEN NAME MIDDLE NAME SURNAME/LAST NAME

HOME ADDRESS

CITY STATE/PROVINCE POSTAL CODE COUNTRY

2.

Are you now or were you ever a member of IEEE? ☐ **Yes** ☐ **No**
If yes, please provide, if known:

MEMBERSHIP NUMBER______________

Grade__________**Year Membership Expired:**__________

3. BUSINESS/PROFESSIONAL INFORMATION

Company Name

Department/Division

Title/Position Years in Current Position

Years in the Profession Since Graduation ☐ PE State/Province

Street Address

City State/Province Postal Code Country

4. EDUCATION

A baccalaureate degree from an IEEE recognized educational program assures assignment of "Member" grade. For others, additional information and references may be necessary for grade assignment.

A. ______________
Baccalaureate Degree Received Program/Course of Study

College/University Campus

State/Province Country Mo./Yr. Degree Received

B. ______________
Highest Technical Degree Received Program/Course of Study

College/University Campus

State/Province Country Mo./Yr. Degree Received

5.

Full signature of applicant

6. DEMOGRAPHIC INFORMATION – ALL APPLICANTS -

Date Of Birth ______________ ☐ Male ☐ Female
Day Month Year

7. CONTACT INFORMATION

Office Phone/Office Fax Home Phone/Home Fax

Office E-Mail Home E-Mail

8. 2016 IEEE AFFILIATE RATES

IEEE DUES Residence	16 Aug 14-28 Feb 15 Pay Full Year	1 Mar -15 Aug 15 Pay Half Year**
United States	$73.50 ☐	$36.75 ☐
Canada (incl. GST)	$77.18 ☐	$38.59 ☐
Canada (incl. HST for PEI)	$83.79 ☐	$41.89 ☐
Canada (incl. HST for Nova Scotia)	$84.58 ☐	$42.26 ☐
Canada (incl. HST for NB, NF and ON)	$83.06 ☐	$41.53 ☐
Canada (incl. GST and QST Quebec)	$84.51 ☐	$42.25 ☐
Africa, Europe, Middle East	$73.50 ☐	$36.75 ☐
Latin America	$73.50 ☐	$36.75 ☐
Asia, Pacific	$73.50 ☐	$36.75 ☐

Canadian Taxes (GST/HST): All supplies, which include dues, Society membership fees, online products and publications (except CD-ROM and DVD media), shipped to locations within Canada are subject to the GST of 5% or the HST of 13%, 14% or 15%, depending on the Province to which the materials are shipped. GST and HST do not apply to Regional Assessments. (IEEE Canadian Business Number 12563 4188 RT0001)
Value Added Tax (VAT) in the European Union: In accordance with the European Union Council Directives 2002/38/EC and 77/388/EEC amended by Council Regulation (EC)792/2002, IEEE is required to charge and collect VAT on electronic/digitized products sold to private consumers that reside in the European Union. The VAT rate applied is the EU member country standard rate where the consumer is resident. (IEEE's VAT registration number is EU826000081)
U.S. Sales Taxes: Please add applicable state and local sales and use tax on orders shipped to **Alabama, Arizona, California, Colorado, District of Columbia, Florida, Georgia, Illinois, Indiana, Kentucky, Massachusetts, Maryland, Michigan, Minnesota, Missouri, New Jersey, New Mexico, New York, North Carolina, Ohio, Oklahoma, West Virginia, Wisconsin.** Customers claiming a tax exemption must include an appropriate and properly completed tax-exemption certificate with their first order.

2016 SPS MEMBER RATES

		16 Aug-28 Feb Pay Full Year	1 Mar-15 Aug Pay Half Year
Signal Processing Society Membership Fee*		$ 22.00 ☐	$ 11.00 ☐

Fee includes: **IEEE Signal Processing Magazine** (electronic and digital), **Inside Signal Proc. eNewsletter** (electronic) and **IEEE Signal Processing Society Content Gazette** (electronic).

Add $17 to enhance SPS Membership and also receive:		$17.00 ☐	$ 8.50 ☐

IEEE Signal Processing Magazine (print) and **SPS Digital Library**: online access to Signal Processing Magazine, Signal Processing Letters, Journal of Selected Topics in Signal Processing, Trans. on Audio, Speech, and Language Processing, Trans. on Image Processing, Trans. on Information Forensics and Security and Trans. on Signal Processing.

Publications available only with SPS membership:

Publication	Format	Full Year	Half Year
Signal Processing, IEEE Transactions on:	Print	$209.00 ☐	$104.50 ☐
Audio, Speech, and Lang. Proc., IEEE/ACM Trans. on:	Print	$160.00 ☐	$ 80.00 ☐
Image Processing, IEEE Transactions on:	Print	$207.00 ☐	$103.50 ☐
Information Forensics and Security, IEEE Trans. on:	Print	$179.00 ☐	$ 89.50 ☐
IEEE Journal of Selected Topics in Signal Processing:	Print	$176.00 ☐	$ 88.00 ☐
Affective Computing, IEEE Transactions on:	Electronic	$ 36.00 ☐	$ 18.00 ☐
Biomedical and Health Informatics, IEEE Journal of:	Print	$ 55.00 ☐	$ 27.50 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$ 65.00 ☐	$ 32.50 ☐
IEEE Cloud Computing	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
IEEE Trans. on Cognitive Comm. &Networking	Electronic	$ 27.00 ☐	$ 13.50 ☐
IEEE Trans. on Computational Imaging	Electronic	$ 28.00 ☐	$ 14.00 ☐
IEEE Trans. on Big Data	Electronic	$ 26.00 ☐	$ 13.00 ☐
IEEE Trans. on Molecular, Biological, & Multi-scale Communications	Electronic	$ 25.00 ☐	$ 12.50 ☐
IEEE Internet of Things Journal	Electronic	$ 26.00 ☐	$ 13.00 ☐
IEEE Trans. on Cloud Computing	Electronic	$ 43.00 ☐	$ 21.50 ☐
IEEE Trans. on Computational Social Systems	Electronic	$ 30.00 ☐	$ 15.00 ☐
IEEE Trans. on Signal & Info Proc. Over Networks	Electronic	$ 28.00 ☐	$ 14.00 ☐
IEEE Biometrics Compendium:	Online	$ 30.00 ☐	$ 15.00 ☐
Computing in Science & Engrg. Mag.:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$ 69.00 ☐	$ 34.50 ☐
Medical Imaging, IEEE Transactions on:	Print	$ 74.00 ☐	$ 37.00 ☐
	Electronic	$ 53.00 ☐	$ 26.50 ☐
	Print & Electronic	$ 89.00 ☐	$ 44.50 ☐
Mobile Computing, IEEE Transactions on:	ELE/Print Abstract/CD-ROM	$ 41.00 ☐	$ 20.50 ☐
Multimedia, IEEE Transactions on:	Electronic	$ 43.00 ☐	$ 21.50 ☐
IEEE MultiMedia Magazine:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$ 69.00 ☐	$ 34.50 ☐
Network Science and Engrg., IEEE Trans. on:	Electronic	$ 34.00 ☐	$ 17.00 ☐
IEEE Reviews in Biomedical Engineering:	Print	$ 25.00 ☐	$ 12.50 ☐
	Electronic	$ 25.00 ☐	$ 12.50 ☐
	Print & Electronic	$ 40.00 ☐	$ 20.00 ☐
IEEE Security and Privacy Magazine:	Electronic and Digital	$ 39.00 ☐	$ 19.50 ☐
	Print	$ 69.00 ☐	$ 34.50 ☐
IEEE Sensors Journal:	Electronic	$ 40.00 ☐	$ 20.00 ☐
Smart Grid, IEEE Transactions on:	Print	$100.00 ☐	$ 50.00 ☐
	Electronic	$ 40.00 ☐	$ 20.00 ☐
	Print & Electronic	$120.00 ☐	$ 60.00 ☐
Wireless Communications, IEEE Transactions on:	Print	$127.00 ☐	$ 63.50 ☐
	Electronic	$ 49.00 ☐	$ 24.50 ☐
	Print & Electronic	$127.00 ☐	$ 63.50 ☐
IEEE Wireless Communications Letters:	Electronic	$ 19.00 ☐	$ 9.50 ☐

*IEEE membership required or requested
Affiliate application to join SP Society only. Amount Paid $______________

9.

IEEE Membership Affiliate Fee (See pricing in Section 8) $______________

Signal Processing Society Fees $______________

Canadian residents pay 5% GST or 13% HST
Reg. No. 125634188 on Society payment(s) & pubs only Tax $______________

AMOUNT PAID WITH APPLICATION TOTAL $______________
Prices subject to change without notice.
☐ **Check or money order enclosed Payable to IEEE on a U.S. Bank**
☐ **American Express** ☐ **VISA** ☐ **MasterCard**
☐ **Diners Club**

Exp. Date/ Mo./Yr.

Cardholder Zip Code Billing Statement Address/USA Only

Full signature of applicant using credit card Date

10. WERE YOU REFERRED?

☐ Yes ☐ No If yes, please provide the follow information:
Member Recruiter Name:______________
IEEE Recruiter's Member Number (Required): ______________

IEEE Signal Processing Society

IEEE