Advanced Signal Processing Workshop

April 16-18th, 2015 – G3, Electrical Engineering Building

Never Stand Still | Faculty of Engineering | School of Electrical Engineering and Telecommunications

This workshop brings together fifteen world-leading researchers in signal processing. It focuses in particular on three broad topics that are of very high current interest:

- Advanced signal processing algorithms that are designed to support higher speed, increased energy efficiency and lower cost wireless communications.
- Statistical signal processing, and in particular signal detection and parameter estimation, using compressed sensing and sparse signal processing
- Automatic processing of speech, including recognition of emotion and mental state from speech.

We gratefully acknowledge funding support for this workshop from UNSW Engineering

The workshop at a glance:

Location: Room G3, Electrical Engineering Building (Campus Map, Google Maps)
Organisers: A/Prof Wei Zhang, Dr Elias Aboutanios, A/Prof Julien Epps and Prof Ambikairajah
Contacts: w.zhang@unsw.edu.au 9385 4033; elias@unsw.edu.au 9385 5010; j.epps@unsw.edu.au 9385 6579
Please register here (registration is free): http://www2.ee.unsw.edu.au/ASPWR_Workshop.html

Program Overview:

	Morning Session	Afternoon Session
Thursday	***Compressive Sensing and Statistical Signal Processing***	***Compressive Sensing and Statistical Signal Processing***
Friday	***Compressive Sensing and Statistical Signal Processing***	***Paralinguistic Speech Processing***
Saturday	***Speech and Signal Processing for Communications***	***Speech and Signal Processing for Communications***

Detailed Program

Thursday April 16th

Morning Session: Compressive Sensing and Statistical Signal Processing

IEEE Distinguished Lecture

11:00 – 12:30 Compressive Covariance Sensing for Enhancing Access to Wireless Spectrum
Prof Zhi Tian, *George Mason University, USA*

Afternoon Session: Compressive Sensing and Statistical Signal Processing

2:00 – 3:00 Randomized Robust Linear Regression for big data applications
Dr Yannis Kopsinis, Marie Curie IEF Fellow, *University of Athens, Greece*

Friday April 17th

9:45 – 10:00	Welcome Dr Elias Aboutanios
	Morning Session: Compressive Sensing and Statistical Signal Processing
10:00 – 10:45	A Sparsity-Perspective to Time-Frequency Signal Representation Prof Moeness Amin, *Villanova University*
10:45 – 11:30	Through-the-Wall Radar Imaging using Joint Bayesian Compressive Sensing Prof Salim Bouzerdoum, *University of Wollongong*
11:30 – 11:45	Coffee break
11:45 – 12:30	Tour of Sparsity-Aware Learning Calling at: Online, Distributed, Robust And Dictionary Learning Prof Sergios Theodoridis, *University of Athens*
12:30 – 1:15	Sparse Signal Recovery and Its Applications Prof Jian Li, *University of Florida*
1:15 – 2:30	Lunch
	Afternoon Session: Paralinguistic Speech Processing
2:30 – 3:15	Feature normalization and model adaptation for robust speech emotion recognition in mismatched conditions Prof Carlos Busso, *University of Texas, Dallas*
3:15 – 4:00	Modeling communication dynamics: from distress indicators to virtual interviewers A/Prof Stefan Scherer, *University of Southern California Institute for Creative Technologies*
4:00 – 4:45	TBA Prof Roland Goecke, *University of Canberra and Australian National University*
5:15	Adjourn
7:00 –	Workshop Dinner (Invited speakers and organisers only)

Saturday April 18th

Speech and Signal Processing for Communications

9:45 – 10:00 Welcome
A/Prof Wei Zhang

10:00 – 10:45 Mispronunciation Detection, Diagnosis and Correct Feedback Generation using Speech Technologies to support Computer-Aided Pronunciation Training for Chinese Learners
Prof Helen Meng, *Chinese University of Hong Kong*

10:45 – 11:30 Search for the “Elementary Particles” in Human Speech – Can we render a monolingual speaker’s speech in a different language?
Prof Frank Soong, *Microsoft Research Asia*

11:30 – 11:45 Coffee break

11:45 – 12:30 Dealing with Imperfections in Human Speech Communication
Prof PC Ching, *Chinese University of Hong Kong*

12:30 – 1:15 How to produce a turbo equalization system with complexity O(logL)?
Prof David Huang, *University of Western Australia*

1:15 – 2:30 Lunch

2:30 – 3:15 Polar Codes: Its Application in HARQ for Wireless Communications and Joint Detection-Decoding Receiver based on Partial Channel State Information
Prof Zhi Ding, *Shanghai Tech University, China*

3:15 – 4:00 Spatial Reuse Precoding for Scalable Downlink Networks
Prof Tim Davidson, *McMaster University, Canada*

4:00 – 4:45 Discussion

4:45 Adjourn

7:00 – Workshop Dinner
(Invited speakers and organisers only)

Speaker Abstracts and Biographies

Compressive Covariance Sensing for Enhancing Access to Wireless Spectrum

Prof Zhi Tian, *George Mason University, USA*

Abstract: Compressive sensing is one of the recent eminent advances in signal processing and statistical learning, with impact to various applications including data sciences, communications, sensor networks, bioinformatics, and medical imaging. It requires information-bearing signals to be sparse over known domains, either naturally or by design. In this talk, I will introduce the fresh notion of compressive covariance sensing, and advocate its exciting implications for (cyclo) stationary processes characterized by second-order statistical descriptors. Such descriptors include (periodic) covariances or frequency, cyclic, angular and Doppler spectra, which already effect signal compression even in the absence of sparsity. Using this key observation, we will demonstrate how the attribute of sparsity can be leveraged more effectively, or, even bypassed when recovering the second-order statistical information of interest. As a leitmotif, we will use the task of wideband spectrum sensing for wireless cognitive radio, which is instrumental for realizing the goal of enhancing access to the radio spectrum. We will present a cyclic feature based compressive spectrum sensing approach for wideband cognitive radios. Using the new framework of compressive covariance sensing, wideband weak signals can be sensed reliably from sub-Nyquist-rate samples in the presence of noise uncertainty, even for (non-sparse) crowded spectrum.

Biography: Dr. Zhi (Gerry) Tian is a Professor in the Electrical and Computer Engineering Department of George Mason University, Fairfax, VA, as of January 2015. Prior to that, she was on the faculty of Michigan Technological University from 2000 to 2014. She served as a Program Director in the Division of Electrical, Communications and Cyber Systems at the US National Science Foundation from 2012 to 2014. Her research interests lie in wireless communications, wireless sensor networks and statistical signal processing. She is an IEEE Fellow. She is an elected member of the IEEE Signal Processing for Communications and Networking Technical Committee (SPCOM‐TC) and a member of the Big Data Special Interest Group IEEE Signal Processing Society. She served as Associate Editor for IEEE Transactions on Wireless Communications and IEEE Transactions on Signal Processing.

Randomized Robust Linear Regression for big data applications

Dr Yannis Kopsinis, Marie Curie IEF Fellow, *University of Athens, Greece*

Abstract: A promising approach when dealing with massive data sets is to apply randomized dimensionality reduction and then operate in lower dimensions. In this talk, the major techniques of randomized linear regression for big data will be presented. Moreover, the case where the available data are sporadically corrupted, a common scenario in big data applications, will be discussed and low computational complexity solutions will be proposed.

Biography: Dr. Yannis Kopsinis is currently holding a Marie Curie IEF fellowship in the Dept. of Informatics and Telecommunications, Univ. of Athens. From Jan. 2004 to Dec. 2005 he has been a research fellow with the Institute for Digital Communications, School of Engineering and Electronics, the University of Edinburgh, UK. From Jan. 2006 to Sep. 2009 he was a senior researcher in the same department. From Jan. 2012 to April 2013 he was a Ramón y Cajal Fellow, University of Granada, Department of Applied Physics, Granada, Spain. He has published more than 50 papers in technical journals and conferences and he has co-authored 3 book chapters. He is also an editorial board member of IET Signal Processing, a member of IEEE and he has served as a member of technical committees of several international conferences. His research interests include adaptive filtering, dimensionality reduction and blind source separation.

A Sparsity-Perspective to Time-Frequency Signal Representation

Prof Moeness Amin, *Villanova University*

Abstract: The talk considers nonstationary signal analysis in view of the signal sparsity properties. We examine these signals, which arise in numerous applications, within the framework of compressive sensing (CS) and sparse reconstructions. We present two general approaches to incorporate sparsity into time-frequency signal presentation (TFSR). In the first approach, quadratic TF distributions (QTFDs) are derived based on optimal multi-task kernel design. In this case, sparseness in the TF domain presents itself as a new design task, adding to those related auto-term preservation and cross-term suppression. In contrast to QTFDs, we also provide a second approach for signal TF signature estimation where sparse reconstruction is used in lieu of direct Fourier transform that maps the signal from the time or one joint-variable domain to another. It is shown that multiple measurement vector methods and block sparsity techniques play a clear and fundamental role in improving signal local power representations. Examples of both approaches are provided. Analysis is supported by simulations with synthesized data, experiments with real Doppler and microDoppler data measurements of radar returns associated with human motions, and by electromagnetic modeling.

Biography: Dr. Amin is the Director of the Center for Advanced Communications, Villanova University, Pennsylvania, USA. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE); Fellow of the European Association for Signal Processing (EURASIP); Fellow of the International Society of Optical Engineering; and a Fellow of the Institute of Engineering and Technology (IET). Dr. Amin is the Recipient of the 2014 IEEE Signal Processing Society Technical Achievement Award; Recipient of the 2009 Individual Technical Achievement Award from EURASIP; Recipient of the 2010 NATO Scientific Achievement Award; Recipient of the 2010 Chief of Naval Research Challenge Award; Recipient of 1997 Villanova University Outstanding Faculty Research Award; and the

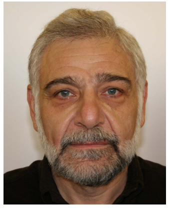
Recipient of the 1997 IEEE Philadelphia Section Award. He is a Recipient of the IEEE Third Millennium Medal, and was a Distinguished Lecturer of the IEEE Signal Processing Society, 2003-2004. Dr. Amin is currently the Chair of the Electrical Cluster of the Franklin Institute Committee on Science and the Arts. He has over 700 journal and conference publications in the areas of Wireless Communications, Time-Frequency Analysis, Sensor Array Processing, Waveform Design and Diversity, Interference Cancellation in Broadband Communication Platforms, Satellite Navigations, Target Localization and Tracking, Direction Finding, Channel Diversity and Equalization, Ultrasound Imaging and Radar Signal Processing. He co-authored 18 book chapters. He is the Editor of the two books "Through the Wall Radar Imaging" and "Compressive Sensing for Urban Radar," published by CRC Press in 2011 and 2014, respectively.

Through-the-Wall Radar Imaging using Joint Bayesian Compressive Sensing

Prof Salim Bouzerdoum, *University of Wollongong*

Abstract: One major challenge in urban sensing applications is to detect stationary targets behind walls and inside enclosed structures. This work addresses the challenging problem of indoor scene reconstruction using through-the-wall radar imaging (TWRI) with only a small subset of frequency measurements. Although not all same frequency measurements are available at each antenna location, the proposed approach reconstructs the antenna signals simultaneously by exploiting sparsity and correlation among different antenna signals. A joint Bayesian sparse model is employed to estimate the antenna signals and to reconstruct the image of the scene. For scene reconstruction, a compact signal model is developed, whereby both the measurement vector and the dictionary are compressed, leading to a more efficient Bayesian sparse scene reconstruction. Furthermore, a subspace-projection technique is applied to the recovered signals to suppress wall clutter and enhance image quality and target detection. The performance is evaluated with simulated and real data. The experimental results show that the proposed approach yields significantly higher signal reconstruction accuracy and requires far fewer measurements for target detection and localization than does the conventional compressed sensing TWRI model.

Biography: Dr. Salim Bouzerdoum (M'89–SM'03) received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle, USA. He is currently Senior Professor of Computer Engineering at the University of Wollongong (UOW), NSW, Australia. From July 1991 to January 1998, he was with the Department of Electrical & Electronic Engineering, Adelaide University, and in 1998 he joined Edith Cowan University as an Associate Professor. In 2004 he was appointed Professor of Computer Engineering and Head of School of Electrical, Computer & Telecom. Engineering at UOW. From 2007 to 2013 he was Associate Dean of Research at UOW. He served on the Australian Research Council College of Experts from 2009 to 2011 and was the Deputy Chair of the EMI panel from 2010 to 2011. Dr. Bouzerdoum is the recipient of the Eureka Prize for Outstanding Science in Support of Defence or National Security in 2011, the IEEE Trans. Consumer Electronics Chester Sall Award in 2005, and a Distinguished Researcher Award (Chercheur de Haut Niveau) from the French Ministry in 2001. He has published over 300 technical articles and graduated 36 Ph.D. and Research Masters students. He served as Associate Editor for 4 International journals, including IEEE TRANS. SYSTEMS, MAN, AND CYBERNETICS (1999–2006). His research interests include image processing, vision, machine learning, and pattern recognition.



Tour of Sparsity-Aware Learning Calling at: Online, Distributed, Robust And Dictionary Learning

Prof Sergios Theodoridis, *University of Athens*

Abstract: Learning sparse models has been a topic at the forefront of research for the last ten years or so. Considerable effort has been invested in developing efficient schemes for the recovery of sparse signal/parameter vectors. Moreover, concepts that have originally been developed around the regression task have been extended to more general and difficult problems, such as low-rank matrix factorization for dimensionality reduction, robust learning in the presence of outliers, dictionary learning for "data-dependent" signal representation. Furthermore, online techniques for sparse modeling estimation are attracting an increasing interest, especially in the context of big data applications. Another area which is gaining in importance is distributed learning over graphs. An area, which was mainly inspired and born within the sensor network discipline, but now lends itself, nicely, for big data processing. In this talk, I touch upon all the previously mentioned problems. Sparse modeling of regression tasks is viewed in its online estimation facet, via convex analytic arguments, based on the set-theoretic framework; the emphasis is on very recent extensions of the theory to include non-convex related constraints, which impose sparsity on the model in a much more aggressive manner compared to the more standard, convex, __-norm related arguments. In spite of the involved non-convexity, still complexity per time iteration exhibits a linear dependence on the number of unknown parameters; furthermore, strong theoretical convergence results have been established. In the sequel, distributed learning techniques are reviewed with an emphasis on greedy-type batch as well as online versions. The task of robust learning in the presence of outliers is then reviewed and new methods, based on the explicit modeling of the outliers, in the context of sparsity-aware learning, will be presented. The new method, based on greedy-type arguments, enjoys a number of merits, compared to more classical techniques. Furthermore, strong theoretical results have been established, for the first time, in such a type of treatment of the robust estimation task. Finally, dictionary learning, in its very recent online and distributed processing framework, is discussed and new experimental as well as theoretical results will be presented.

Biography: Sergios Theodoridis is currently Professor of Signal Processing and Machine Learning in the Department of Informatics and Telecommunications of the University of Athens. His research interests lie in the areas of Adaptive Algorithms, Distributed and Sparsity-Aware Learning, Machine Learning and Pattern Recognition, Signal Processing for Audio Processing and Retrieval. He is the author of the book "Machine Learning: A Bayesian and Optimization Perspective" Academic Press (1030 pages), 2015, the co-author (with K. Koutroumbas) of the best-selling book "Pattern Recognition", Academic Press (960 pages), 4th ed. 2009, the co-author of the book "Introduction to Pattern Recognition: A MATLAB Approach", Academic Press, 2010, the coeditor (with N. Kaloupstidis) of the book "Efficient Algorithms for Signal Processing and System Identification", Prentice Hall 1993, and the co-author of three books in Greek, two of them for the Greek Open University. He currently serves as Editor-in-Chief for the IEEE Transactions on Signal Processing. He is Editor-in-Chief for the Signal Processing Book Series, Academic Press and co-Editor in Chief (with Rama Chellapa) for the E-Reference Signal Processing, Elsevier. He is the co-author of seven papers that have received Best Paper Awards including the 2014 IEEE Signal Processing Magazine best paper award and the 2009 IEEE Computational Intelligence Society Transactions on Neural Networks Outstanding Paper Award. He is the recipient of the 2014 IEEE Signal Processing Society Education Award and the 2014 EURASIP Meritorious Service Award. He has served as an IEEE Signal Processing Society Distinguished Lecturer. He was Otto Monstead Guest Professor, Technical University of Denmark, 2012, and holder of the Excellence Chair, Dept. of Signal Processing and Communications, University Carlos III, Madrid, Spain, 2011. He serves (2014-2016) as Distinguished Lecturer for the IEEE Circuits and Systems Society. He was the general chairman of EUSIPCO-98, the Technical Program co-chair for ISCAS-2006 and ISCAS-2013, co-chairman and co-founder of CIP-2008, co-chairman of CIP-2010 and Technical Program co-chair of ISCCSP-2014. He has served as President of the European Association for Signal Processing (EURASIP), as a member of the Board of Governors for the IEEE CAS Society, as a member of the Board of Governors (Member-at-Large) of the IEEE SP Society and as a Chair of the Signal Processing Theory and Methods (SPTM) technical committee of IEEE SPS. He has served as a member of the Greek National Council for Research and Technology and he was Chairman of the SP advisory committee for the Edinburgh Research Partnership (ERP). He has served as vice chairman of the Greek Pedagogical Institute and he was for four years member of the Board of Directors of COSMOTE (the Greek mobile phone operating company). He is Fellow of IET, a Corresponding Fellow of the Royal Society of Edinburgh (RSE), a Fellow of EURASIP and a Fellow of IEEE.

Sparse Signal Recovery and Its Applications

Prof Jian Li, *University of Florida*

Abstract: We consider nonparametric adaptive spectral analysis of complex-valued data sequences with possibly missing samples occurring in arbitrary patterns. We will present two high-resolution spectral estimation algorithms: the Iterative Adaptive Approach (IAA) and the Sparse Learning via Iterative Minimization (SLIM) method. Both algorithms are easy to use in practical applications since they are user parameter free. Both can be used to significantly improve the spectral estimation performance, including enhanced resolution and reduced sidelobe levels. Moreover, we have considered fast implementations of these algorithms using the Conjugate Gradient (CG) technique and the Gohberg-Semencul-type (GS) formula. Our proposed implementations fully exploit the structure of the steering matrices and maximize the usage of the Fast Fourier Transform (FFT), resulting in much lower computational complexities as well as much reduced memory requirements. The effectiveness of the adaptive spectral estimation algorithms is demonstrated via several numerical examples including both 1-D spectral estimation and 2-D synthetic aperture radar imaging and automatic target recognition examples.

Biography: Jian Li (PhD, Ohio State University) teaches signal processing at the University of Florida. Dr. Li's publications include Robust Adaptive Beamforming (2005, Wiley), Spectral Analysis: the Missing Data Case (2005, Morgan & Claypool), MIMO Radar Signal Processing (2009, Wiley), and Waveform Design for Active Sensing Systems -- A computational approach (2011, Cambridge University Press). Her research interests include spectral estimation, statistical and array signal processing, and their applications to radar, sonar, communications, and medical imaging.

Feature normalization and model adaptation for robust speech emotion recognition in mismatched conditions

Prof Carlos Busso, *University of Texas, Dallas*

Abstract: While early studies on speech emotion recognition have reported classification performance between 50%-85%, the accuracy of these systems significantly drops when models are tested on natural unconstrained settings. The mismatch between train and test conditions are due to differences between read and spontaneous speech, prototypical versus subtle emotions, and idiosyncratic and heterogeneity differences in the externalization of emotions, and myriad contextual influences. The lack of generalization of speech emotion recognition models to recognize expressive behavior during natural, unrestricted human interaction is the key barrier to deploying affective-aware technology to real-world applications, including speech summarization. We present feature normalization and model adaptation methods to overcome these limitations by exploring algorithms rooted in the externalization process of verbal and nonverbal behaviors, representing a transformative breakthrough in the area of affective behavioral analysis. The algorithmic solutions rely on the latest developments in transfer learning and feature normalization. By integrating novel theoretical and experimental methods with realistic applications, our research aims to enable systematic algorithm design and its robustness to real world non-prototypical expressive speech.

Biography: Carlos Busso is an Assistant Professor at the Electrical Engineering Department of The University of Texas at Dallas (UTD). He received his B.S (2000) and M.S (2003) degrees with high honors in electrical engineering from University of Chile, Santiago, Chile, and his Ph.D (2008) in electrical engineering from University of Southern California (USC), Los Angeles, USA. He was selected by the School of Engineering of Chile as the best Electrical Engineer graduated in 2003 across Chilean universities. At USC, he received a Provost Doctoral Fellowship from 2003 to 2005 and a Fellowship in Digital Scholarship from 2007 to 2008. At UTD, he leads the Multimodal Signal Processing (MSP) laboratory [http://msp.utdallas.edu]. In 2014, he received the ICMI Ten-Year Technical Impact Award. He also received the Hewlett Packard Best Paper Award at the IEEE ICME 2011 (with J. Jain). He is the co-author of the winner paper of the Classifier Sub-Challenge event at the Interspeech 2009 emotion challenge. His research interests are in digital signal processing, speech and video processing, and multimodal interfaces. His current research includes the broad areas of affective computing, multimodal human-machine interfaces, modeling and synthesis of verbal and nonverbal behaviors, sensing human interaction, in-vehicle active safety system, and machine learning methods for multimodal processing.

Modeling communication dynamics: from distress indicators to virtual interviewers

A/Prof Stefan Scherer, *University of Southern California Institute for Creative Technologies*

Abstract: Unlike laboratory values or radiographic images, clinical material from a mental health encounter consists of complex human interactions between patient and provider characterizing mood, affect, and behavior. Because much of this information is subjective and difficult or impossible to annotate while maintaining affordability of clinical care, it becomes inaccessible to treatment planning. Recent progress in facial feature tracking, voice analysis, and body tracking has opened the door to new applications for automatic nonverbal behavior analysis and enables us to derive behavioral markers or descriptors from mental health encounters. Our research on automatic behavior descriptors is specifically tailored to provide additional quantitative information to the healthcare providers. It encompasses recent advances in machine learning, speech signal processing, and computer vision, and includes modeling of multimodal and interpersonal dynamics observed between healthcare providers and patients.

Based on the derived communication dynamics models, we propose the use of virtual interviewers, that incorporate and automatically perceive the patients' behavior. Virtual human interviewers hold several advantages over their natural counterparts: their behavior can be pre-programmed to the slightest detail and increases the available level of control for the investigators or clinical personnel over the assessment process and the presentation of stimuli. In addition, virtual human interviewers may allow researchers to obtain more, or richer samples of data than with real human interviewers. In fact, interacting with a virtual human can increase participants' willingness to say more. In particular, an investigation of the effects of framing the character as human-controlled or autonomous showed that participants felt more comfortable disclosing personal information with a character that was framed as autonomous than when it was framed as human-controlled.

Biography: In July of 2011, Dr. Stefan Scherer finished his doctoral studies to receive the degree of Dr. rer. nat. from the faculty of Engineering and Computer Science at Ulm University. Scherer's thesis, entitled "Analyzing the User's State in HCI: From Crisp Emotions to Conversational Dispositions," received the grade summa cum laude (i.e., with distinction). As a postdoctoral researcher, he had the opportunity to collect experience at the internationally renowned Trinity College Dublin in Ireland as well as at the University of Southern California Institute for Creative Technologies. Currently, Scherer is working as Research Assistant Professor in the Computer Science Department and as Research Associate at the University of Southern California Institute for Creative Technologies. His research fields of interest are human machine interaction, social and speech signal processing, machine learning, and affective computing.

Title

Prof Roland Goecke, *University of Canberra and Australian National University*

Abstract: TBA

Biography: Roland Goecke is a Professor of Affective Computing at the University of Canberra. He is the Head of the Vision and Sensing Group and Director of the Human-Centred Technology Research Centre. He received his Masters degree in Computer Science from the University of Rostock, Germany, in 1998 and his PhD in Computer Science from the Australian National University, Canberra, Australia, in 2004.

Mispronunciation Detection, Diagnosis and Correct Feedback Generation using Speech Technologies to support Computer-Aided Pronunciation Training for Chinese Learners

Prof Helen Meng, *Chinese University of Hong Kong*

Abstract: This talk presents an ongoing research initiative in the development of speech technologies that strives to raise the efficacy of computer-aided pronunciation training, especially for Chinese learners of English. Our approach is grounded on the theory of language transfer and involves a systematic phonological comparison between the primary language (L1 being Chinese) and secondary language (L2 being English) to predict possible realizations that constitute mispronunciations in L2 English. Mispronunciation diagnosis aims to support the design of pedagogical and remedial instructions, which involves audiovisual text-to-speech synthesis technologies for corrective feedback generation. This talk offers an overview of the technologies, related experimental results and ongoing work as well as future plans.

Biography: Helen Meng received all her degrees (S.B., S.M. and Ph.D.) in electrical engineering from MIT. She is Professor and Chairman of the Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong. She is also the Founding Director of the Microsoft-CUHK Joint Laboratory for Human-Centric Computing and Interface Technologies, which has been recognized as a Ministry of Education of China (MoE) Key Laboratory since 2008. In 2013, she helped establish the CUHK Big Data Decision Analytics Research Center and serves as its Founding Director. She also served as Associate Dean (Research) of the Faculty of Engineering from 2006 to 2010. She is the former Editor-in-Chief of the IEEE Transactions on Audio, Speech and Language Processing, a member of the IEEE Speech and Language Processing Technical Committee (for a second term), and is elected into the IEEE Board of Governors in 2014. Helen's professional services include membership in the Steering Committee on eHealth Record Sharing of the HKSAR Government, Council membership of Hong Kong Productivity Council and Open University of Hong Kong, as well as former memberships in the HKSAR Government's Digital 21 Strategy Advisory Committee and Research Grants Council. Helen has also served in review panels of various funding agencies, including the Hong Kong SAR Government's Innovation and Technology Commission, Swedish Research Council European Research Infrastructure Initiative, and the National Centres of Competence in Research of the Swiss National Science Foundation. Professor Meng received the MoE Higher Education Outstanding Scientific Research Output Award in Technological Advancements in 2009. In previous years, she has also received the Exemplary Teaching Award, Young Researcher Award, as well as Service Award from the CUHK Faculty of Engineering. Helen was elected one of the inaugural Distinguished Lecturers of the APSIPA in 2012. She is also a Fellow of the Hong Kong Computer Society, Hong Kong Institute of Engineers and also a Fellow of IEEE for contributions to spoken language and multimodal systems.

Dealing with Imperfections in Human Speech Communication

Prof Pak-Chung Ching, *Chinese University of Hong Kong*

Abstract: Human speech communication involves a series of complicated processes that require inter-disciplinary knowledge to understand and comprehend thoroughly. From signal processing points of view, we are interested in the unique properties of acoustic speech signals and how these properties contribute to effective communication from the speaker to the listener. Along the speech communication pathway, there are many imperfections that are related to disabilities of the speakers or listeners. Regardless of the causes of the disabilities, the acoustic speech signals transmitted from the speaker to the listener remain the most relevant physical observables, which can be analyzed, intervened or manipulated by signal processing techniques. In this talk, we will describe our recent efforts on applying speech processing algorithms and spoken language technologies to address various problems related to speech and hearing disorders. They include signal enhancement for improving tone perception of cochlear implant users, pitch control for electro-larynx systems, objective assessment and analysis of disordered voice and aphasia speech.

This is joint work with Prof. Tan Lee at CUHK.

Biography: Pak-chung Ching received the B. Eng. (1st Class Honors) and Ph.D. degrees from the University of Liverpool, UK, in 1977 and 1981 respectively. From 1981 to 1982 he was Research Officer at the University of Bath, UK. In 1982, Prof. Ching returned to Hong Kong and joined the then Hong Kong Polytechnic as a lecturer. Since 1984, he has been with the Department of Electronic Engineering of the Chinese University of Hong Kong (CUHK), where he is currently Choh-Ming Li Professor of Electronic Engineering. He was Department Chairman from 1995 to 1997, Dean of Engineering from 1997 to 2003 and Head of Shaw College from 2004 to 2008. He became Director of the Shun Hing Institute of Advanced Engineering in 2004. From 2006 to end of 2014, Prof Ching was appointed as Pro-Vice-Chancellor/Vice-President of CUHK. Between 2013 to 2014, Prof. Ching also took up the Directorship of the CUHK Shenzhen Research Institute. Prof. Ching is very active in promoting professional activities, both in Hong Kong and overseas. He was a council member of the Institution of Electrical Engineers (IEE), past chairman of the IEEE Hong Kong Section, an associate editor of the IEEE Transactions on Signal Processing from 1997 to 2000 and IEEE Signal Processing Letters from 2001 to 2003. He was also a member of the Technical Committee of the IEEE Signal Processing Society from 1996 to 2004. He was appointed Editor-in-Chief of the HKIE Transactions between 2001 and 2004. He has been an Honorary Member of the editorial committee for Journal of Data Acquisition and Processing since 2000. Prof. Ching has been instrumental in organizing many international conferences in Hong Kong including the 1997 IEEE International Symposium on Circuits and Systems where he was the Vice-Chairman, and the 2003 IEEE International Conference on Acoustics, Speech and Signal Processing where he was the Technical Program Chair. He also serves as Technical Program Co-Chair of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing. Prof Ching was awarded the IEEE Third Millennium Award in 2000 and the HKIE Hall of Fame in 2010. In addition, Prof. Ching also plays an active role in community services. In 2010, he was awarded the Bronze Bauhinia Star (BBS) by the HKSAR Government. He is presently Chairman of the Hong Kong Council for Testing and Certification; Chairman of the Veterinary Surgeons Board of Hong Kong; Member of the Steering Committee on Innovation and Technology; and Council Member of the Shaw Prize Foundation. He is also involved in various capacities in the Hong Kong Applied Science and Technology Research Institute Company Ltd, HK R&D Centre for Logistics and Supply Chain Management Enabling Technologies and Nano and Advanced Materials Institute Limited. He was Member of The Greater Pearl River Delta Business Council from 2012 to 2013, Panel Member of the Research Grant Council from 2005 to 2010, Member of the Governing Council of the Hong Kong Quality Assurance Agency from 2005 to 2009, Member of the Copyright Tribunal from 2003 to 2009, Member of Electrical Safety Advisory Committee of the Electrical and Mechanical Services Department from 2002 to 2008, Member of the Consumer Council from 2001 to 2006 and Chairman of the Hong Kong Accreditation Advisory Board from 2000 to 2005. Prof. Ching is also a Guest Professor at Tsinghua University, Beijing University of Posts and Telecommunications, Huazhong University of Science and Technology, Southeast University and Institute of Acoustics, Chinese Academy of Sciences.

Search for the "Elementary Particles" in Human Speech – Can we render a monolingual speaker's speech in a different language?

Prof Frank Soong, *Microsoft Research Asia*

Abstract: In this talk, we will raise a challenging question: Can we find the "elementary particles" of a person's speech in one language and use them for rendering his/her voice in a different language? A positive "yes" answer and the found "elementary particles" can have many useful applications, e.g. mixed code text-to-speech (TTS), language learning, speech-to-speech translation, etc. We will answer the question first on how to train a TTS in a target language (L2) with speech corpus collected from a source speaker in his own mother tongue (L1), which is different from the target language, along with a speech corpus recorded by a reference speaker in L2. We then use the "trajectory tiling algorithm," invented by us for synthesizing high quality, unit selection TTS, to "tile" the trajectories of all sentences in the reference speaker's corpus with the "closest" speech segments in the source speaker's data. To make the tiling proper across two different (reference and source) speakers, the difference between their speech signals needs to be equalized with appropriate vocal tract length normalization, e.g., a bilinear warping function or formant mapping. All tiled sentences are then used to train a new HMM-based TTS of the monolingual speaker but in the reference speaker's language. Different length units of the 'elementary particles" have been tried and a label-less frame length (10 ms) segments have been found to yield the best TTS quality.

In addition to cross-lingual TTS, some experimental results also show that training a speech recognizer with speech data of different languages tends to improve the ASR performance

in each individual language. Also, besides the discovered audio "elementary particles" for cross-lingual TTS, the mouth shapes of a mono-lingual speaker have also been found adequate for rendering the lips movement of talking heads in different languages. Various demos will be shown to illustrate our findings.

Biography: Frank K. Soong, a Principal Researcher and Research Manager, Speech Group, Microsoft Research Asia (MSRA), Beijing, works on fundamental research on speech and its practical applications. His professional research career spans over 30 years, first with Bell Labs, US, then with ATR, Japan, before joining MSRA in 2004. At Bell Labs, he worked on stochastic modeling of speech signals, optimal decoder algorithm, speech analysis and coding, speech and speaker recognition. He was responsible for developing the recognition algorithm which was developed into voice-activated mobile phone products rated by the Mobile Office Magazine (Apr. 1993) as the "outstandingly the best". He is a co-recipient of the Bell Labs President Gold Award for developing the Bell Labs Automatic Speech Recognition (BLASR) software package.

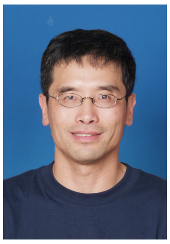
He has served as a member of the Speech and Language Technical Committee, IEEE Signal Processing Society and other society functions, including Associate Editor of the IEEE Speech and Audio Transactions and chairing IEEE Workshop. He published extensively with more than 200 papers and co-edited a widely used reference book, Automatic Speech and Speech Recognition- Advanced Topics, Kluwer, 1996. He is a visiting professor of the Chinese University of Hong Kong (CUHK) and a few other top-rated universities in China. He is also the co-Director of the National MSRA-CUHK Joint Research Lab. He got his BS, MS and PhD from National Taiwan Univ., Univ. of Rhode Island, and Stanford Univ, all in Electrical Eng. He is an IEEE Fellow "for contributions to digital processing of speech".

How to produce a turbo equalization system with complexity O(logL)?

Prof David Huang, *University of Western Australia*

Abstract: Following the great success of turbo codes in 1993, turbo equalization was conceived as an approach to achieving bandwidth efficient digital communications over inter-symbol interference channels. However, for more two decades, the prohibitive complexity of turbo equalization has prevented its application in practice. Powered by probabilistic graphical models and through proper approximations, we have managed to reduce the complexity of turbo equalization, including the Soft-Input and Soft-Output (SISO) equalizer and the channel estimator, to the level of O(logL) per symbol per iteration, where L is the number of channel taps. In this talk, how this has been achieved will be presented and the ramification of this progress will also be discussed.

Biography: Defeng (David) Huang received the B.E.E.E. and M.E.E.E. degree in electronic engineering from Tsinghua University in 1996 and 1999, respectively, and the PhD degree in electrical and electronic engineering from HKUST in 2004. He joined the School of Electrical, Electronic and Computer Engineering at the University of Western Australia (UWA) in 2005 as a lecturer, and is now a professor. Before joining UWA, he was a lecturer at Tsinghua University. Dr. Huang served as an editor for the IEEE transactions on Wireless Communications from 2005 to 2011 and currently serves as an editor for the IEEE Wireless Communications Letters. His research interests mainly focus on digital communications including wireless communications, underwater acoustic communications, and power line communications.



Polar Codes: Its Application in HARQ for Wireless Communications and Joint Detection-Decoding Receiver based on Partial Channel State Information

Prof Zhi Ding, *Shanghai Tech University, China*

Abstract: Based on the concept of channel polarization, polar codes can provably achieve capacity as forward error correction codes. Polar codes can be systematically constructed and facilitate low complexity encoding and decoding. Though specially designed for binary channels, polar codes have been investigated in applications involving more general channels. In this work, we investigate the utilization of polar codes in hybrid ARQ (HARQ) systems. To improve bandwidth efficiency, consider packet retransmissions without explicit training for channel estimations after training for channel estimation in the original packet transmission. In this partial channel state information (CSI) scenario, we develop a joint signal detection and decoding receiver for polar coded HARQ. We show that a joint receiver can effectively utilize a linear programming algorithm to achieve competitive results with relatively low computational complexity. This concept can be further improved and generalized for broader communication systems including relay networks.

Biography: Zhi Ding (S'88-M'90-SM'95-F'03) is a Professor of Electrical and Computer Engineering at the University of California, Davis. He received his Ph.D. degree in Electrical Engineering from Cornell University in 1990. From 1990 to 2000, he was a faculty member of Auburn University and later, University of Iowa. Prof. Ding has held visiting positions in Australian National University, Hong Kong University of Science and Technology, NASA Lewis Research Center and USAF Wright Laboratory. Prof. Ding has active collaboration with researchers from several countries including Australia, China, Japan, Canada, Taiwan, Korea, Singapore, and Hong Kong. Dr. Ding is a Fellow of IEEE and has been an active member of IEEE, serving on technical programs of several workshops and conferences. He was associate editor for IEEE Transactions on Signal Processing from 1994-1997, 2001-2004, and associate editor of IEEE Signal Processing Letters 2002-2005. He was a member of technical committee on Statistical Signal and Array Processing and member of technical committee on Signal Processing for Communications (1994-2003). Dr. Ding was the Technical Program Chair of the 2006 IEEE Globecom. He is also an IEEE Distinguished Lecturer (Circuits and Systems Society, 2004-06, Communications Society, 2008-09). Dr. Ding is a coauthor of the textbook: Modern Digital and Analog Communication Systems, 4th edition, Oxford University Press, 2009. Prof. Ding is currently at the inaugural ShanghaiTech University to play a key role in the foundation of the School of Information Science and Technology.

Spatial Reuse Precoding for Scalable Downlink Networks

Prof Tim Davidson, *McMaster University, Canada*

Abstract: Inter-cell interference in cellular wireless networks has traditionally been managed by "reuse" strategy in which frequency bands (or time slots) are re-used in geographically

separated cells (or fractions thereof). Such interference avoidance approaches greatly simplify the operation of the network, as the operation of the cells is decoupled, but they can be quite inefficient as they leave significant communication resources idle. More recent notions of cooperative transmission, coordinated transmission and interference alignment have shown that when cells are clustered and interference is managed jointly, many of those resources can be activated. However, the size of the cluster is limited by the reliance of most such schemes on knowledge of the "inter-cell" channels and by the computational cost of determining appropriate signalling schemes.

In this talk, we outline the principles of a signally scheme that seeks to enable more degrees of freedom than traditional interference avoidance schemes, while retaining the advantages that inter-cell interference can be cancelled and the network decoupled without requiring inter-cell channel state information at the transmitters. This decoupling facilitates a "reuse" strategy that enables the precoding scheme to be scaled to extended structured networks, and even to certain heterogeneous networks. The signalling scheme is developed for quasi-static downlink channels and exploits the decomposable (Kronecker) structure of the equivalent channel matrix by imposing that structure on a factor of the precoding matrices. The reuse strategy is spatial in both the geographical sense and in the sense that it is based on the fact that a time-invariant channel can change the basis of a signal subspace, but does not change the subspace in which the signal lies. Preliminary simulation results indicate that the proposed scheme can provide substantially higher rates to the cell edge users than schemes that ignore interference, schemes that avoid interference, and schemes based on interference alignment over clusters of small numbers of cells.

This talk presents joint work with Ahmed Medra.

Biography: Tim Davidson received the B.Eng. (Hons. I) degree in electronic engineering from the University of Western Australia (UWA), Perth, in 1991 and the D.Phil. degree in engineering science from the University of Oxford, U.K., in 1995.

He is a Professor in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he is currently serving as Chair of the Department. Previously, he had served as Acting Director of the School of Computational Engineering and Science for two years. His research interests lie in the general areas of communications, signal processing, and control.

Dr. Davidson received the 1991 J. A. Wood Memorial Prize from UWA, the 1991 Rhodes Scholarship for Western Australia, and a 2011 Best Paper Award from the IEEE Signal Processing Society. He has served as an Associate Editor of three IEEE transactions and as a Guest Co-Editor of two issues of IEEE journals. He was a General Co-Chair for the 2014 IEEE International Workshop on Signal Processing Advances in Wireless Communications, and a Technical Program Co-Chair for the 2014 IEEE Global Conference on Signal and Information Processing. Dr. Davidson is currently serving as the Past-Chair of the IEEE Signal Processing Society’s Technical Committee on Signal Processing for Communications and Networking. He is a Registered Professional Engineer in the Province of Ontario.