

Guest Editorial: Machine Learning for Signal Processing

Shigeru Katagiri • Atsushi Nakamura • Tülay Adalı •
Jianhua Tao • Jan Larsen • Tieniu Tan

Published online: 6 February 2014
© Springer Science+Business Media New York 2014

One of the fundamental goals of signal processing is to model observed signals or their underlying processes accurately. This is a very challenging problem due to issues such as observation noise and the finite nature of observations. To overcome these difficulties, signal processing has proactively employed the relatively new and powerful discipline of machine learning. The type of signal processing that incorporates machine learning as its booster, i.e., Machine Learning for Signal Processing (MSLP), is now working as a promising technological paradigm that opens the way to discovering salient features from noisy signals, predicting unseen signals, and achieving accurate modeling of the process that generates signals.

A large number of papers related to MLSP have been published in different places. However, to advance research in the area effectively, one comprehensive forum would clearly be advantageous. Motivated by this need, we organized the publication project of this special issue on MSLP. Twenty-one papers were submitted to the issue, and the eleven papers presented here were selected through a rigorous review process. The

selected papers cover most of the key factors of signal processing, in particular its important sub-field of pattern classification.

Pattern classification basically consists of a feature-extraction (selection) stage, which converts an input pattern to features, and a classification stage, which labels a set of converted features as one of various preset class indexes. Research on pattern classification sometimes focuses on either of these component stages. Each of the selected papers more or less refers to both of the stages, but they can be divided into three groups to give readers a clear understanding of the issue and the relative positions of the included papers. The first four papers focus on theoretical or algorithmic research topics related to the classification stage by studying ways of designing (training) class models. The next three papers study the extraction of salient features, or those robust to unseen patterns, from a limited amount of input data. Finally, the remaining four papers study applications of signal modeling and/or pattern classification to particular types of input data, such as face images and speech signals.

S. Katagiri (✉)
Department of Information System Design, Faculty of Science and Engineering, Doshisha University, 1-3 Tataramiyakodani, Kyotanabe, Kyoto 610-0394, Japan
e-mail: skatagiri@mail.doshisha.ac.jp

A. Nakamura
Media Information Laboratory, NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, 2-4 Hikaridai, Seika-Cho, Soraku-Gun, Kyoto 619-0237, Japan
e-mail: nakamura.atsushi@lab.ntt.co.jp
e-mail: atsushi@ieee.org

T. Adalı
Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, ITE 324, 1000 Hilltop Circle, Baltimore, MD 21250, USA
e-mail: adali@umbc.edu

J. Tao
National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, 95 Zhong Guan Cun East Road, Beijing 100190, China
e-mail: jhtao@nlpr.ia.ac.cn

J. Larsen
Department of Informatics and Mathematical Modelling, Technical University of Denmark, Matematiktorvet, building 303B, Kongens Lyngby 2800, Denmark
e-mail: janla@dtu.dk

T. Tan
National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, 95 Zhong Guan Cun East Road, Beijing 100190, China
e-mail: tnt@nlpr.ia.ac.cn

In contrast to the traditional classification scheme that labels an input as one of preset class indexes, the paper entitled “An Efficient Gradient-based Approach to Optimizing Average Precision through Maximal Figure-of-Merit Learning” discusses an effective training procedure for classifying an input in order of multiple class indexes, or in other words, ranking an input, which plays an important role in document and image retrieval. In “Robust and Efficient Pattern Classification using Large Geometric Margin Minimum Classification Error Training” and “Minimum Classification Error Training Incorporating Automatic Loss Smoothness Determination,” the authors focus on a fundamental scheme of labeling an input to one of several preset classes, and they discuss the recent advances in the popular discriminative training method for pattern classifier design, Minimum Classification Error (MCE) training. Novelty detection from an observed time-series signal can be considered a one-class classification, because it involves only the modeling of normal class for observations. The paper entitled “Extending the Generalised Pareto Distribution for Novelty Detection in High-Dimensional Spaces” elaborates a theoretical aspect of modeling complex and high-dimensional signals, and it experimentally evaluates an extended use of the Generalized Pareto Distribution on patient physiological monitoring data.

When the amount of training data is severely limited, training sometimes adopts the assumption that pattern samples and even class model parameters are based on some distribution function. In “Structural Bayesian Linear Regression for Hidden Markov Models,” the authors propose a variational Bayesian approach to the modeling of linear regression parameters for Hidden Markov Model (HMM)-based speech recognition. In “Predictive Distribution of the Dirichlet Mixture Model by Local Variational Inference Method,” the authors propose an algorithm to calculate the predictive distribution of the Dirichlet Mixture Model with the Local Variational Inference method, and they provide experimental evaluation results on EEG signal data. Then, in “Bayesian Sparse Topic Model,” the authors present a new Bayesian sparse learning approach, based on sparse Latent Dirichlet Allocation, to select salient lexical features for sparse topic modeling.

Seemingly, the higher the dimension of the converted features is, the easier it is to classify the set of those features. However, this is not actually true, due to such causes as sparseness or robustness to unseen signals. To alleviate this problem, in “Face Recognition with Integrating Multiple Cues,” the authors study information-fusion-based techniques that integrate multiple features. On the other hand, the classification or signal processing of face image data must countervail the excessive high-dimensionality of input. To alleviate this computational difficulty, the paper entitled “Uniform Local Derivative Patterns and Their Application in Face Recognition” focuses on the local feature representation power of Local Derivative Patterns (LDPs) and improves the computational

burden of LDPs. In “Symbolic Time Series Analysis of Temporal Gait Dynamics,” the authors study an approach of modeling high-dimensional time-series input with symbol sequences, or in other words, classifying it as symbolic class indexes. The final paper of this special issue is “Pitch-Scaled Spectrum based Excitation Model for HMM-based Speech Synthesis.” This paper does not deal with classification but studies a way of modeling an input speech signal accurately with the aim of improving the quality of HMM-synthesized speech.

This special issue is not extensive enough to introduce the entire frontline of MLSP research, but it efficiently spotlights key advances in the sub-fields of MLSP, that is, pattern classification and face/speech processing. The editors hope that this issue will contribute toward further advances in signal processing technologies.



Shigeru Katagiri received B.E. and M.E. degrees in Electrical Engineering and a Dr. Eng. degree in Information Engineering from Tohoku University, Sendai, Japan, in 1977, 1979, and 1982, respectively. From 1982 to 1986, he worked at the Electrical Communication Laboratories, Nippon Telegraph and Telephone Public Corporation (currently NTT), Tokyo. From 1986 to 1998, he was with the Advanced Telecommunications Research Institute International (ATR), Kyoto. In 1999 he returned to NTT, and until 2006 he was with the NTT Communication Science Laboratories (CS Labs), Kyoto, where he served as several research manager positions including the Director of CS Labs. Since 2006, he has been with Doshisha University, where he is currently a Professor in the Graduate School of Science and Engineering. His research interests include pattern recognition, speech signal processing, and computer-assisted media space for supporting remote collaboration. Dr. Katagiri is an Associate Member of the Science Council of Japan, an IEEE Fellow, and an NTT R&D Fellow.



Atsushi Nakamura received the B.E., M.E., and Dr.Eng. degrees from Kyushu University, Fukuoka, Japan, in 1985, 1987 and 2001, respectively. In 1987, he joined Nippon Telegraph and Telephone Corporation (NTT), where he engaged in the research and development of network service platforms, at Musashino Electrical Communication Laboratories, Tokyo, Japan. From 1994 to 2000, he was with Advanced Telecommunications Research (ATR) Institute, Kyoto, Japan, as a Senior Researcher, working on the research of spontaneous speech recognition and development of speech translation systems. Since April, 2000, he has been with NTT Communication Science

Laboratories, and currently heads a department of signal processing research. Dr. Nakamura is a senior member of IEEE, and serves as the Chair of Kansai Chapter, IEEE Signal Processing Society, as a member of IEEE Machine Learning for Signal Processing Technical Committee.



Tülay Adalı received the Ph.D. degree in Electrical Engineering from North Carolina State University, Raleigh, NC, USA, in 1992 and joined the faculty at the University of Maryland Baltimore County (UMBC), Baltimore, MD, USA, the same year. She is currently a professor in the Department of Computer Science and Electrical Engineering at UMBC. She has been active in conference and workshop organizations. She was the general or technical co-chair of the IEEE

Machine Learning for Signal Processing (MLSP) and Neural Networks for Signal Processing Workshops 2001–2008, and helped organize a number of conferences including the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). She has served or currently serving on numerous editorial boards of journals of primarily the IEEE, and technical committees of the IEEE Signal Processing Society. She was the chair of the technical committee on MLSP, 2003–2005 and 2011–2013. Prof. Adalı is a Fellow of the IEEE and the AIMBE, recipient of a 2010 I.E. Signal Processing Society Best Paper Award, 2013 University System of Maryland Regents' Award for Research, and an NSF CAREER Award. She was an IEEE Signal Processing Society Distinguished Lecturer for 2012 and 2013. Her research interests are in the areas of statistical signal processing, machine learning for signal processing, and biomedical data analysis.



Jianhua Tao (M'98) received the M.S. degree from Nanjing University, Nanjing, China, in 1996 and the Ph.D. degree from Tsinghua University, Beijing, China, in 2001.

He is currently a Professor with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing. His current research interests include speech synthesis and recognition, human–computer interaction, and emotional information processing.

He has published more than 60 papers in major journals and proceedings, such as the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, ICASSP, Interspeech, ICME, ICPR, ICCV, ICIP, etc. In 2006, he was elected as Vice-Chairperson of the ISCA Special Interest Group of Chinese Spoken Language Processing (SIG-CSLP), and Executive Committee member of the HUMAINE association. He is the subject editor for the Speech Communication (SPEECH COMMUN), the Editorial Board Member for the Journal on Multimodal

User Interfaces (JMUI), the International Journal of Synthetic Emotions (IJSE), and the Steering Committee Member for the IEEE TRANSACTIONS ON AFFECTIVE COMPUTING.



Jan Larsen received the M.Sc. and Ph.D. degrees in electrical engineering from the Technical University of Denmark (DTU) in 1989 and 1994. Dr. Larsen is Associate Professor of Digital Signal Processing at Department of Applied Mathematics and Computer Science, DTU. Jan Larsen has authored and co-authored more than 125 papers and book chapters within the areas of nonlinear statistical signal processing, machine learning, neural networks and datamining with applications

to biomedicine, monitoring systems, multimedia, audio, and webmining. According to Google Scholar the no. of citations is 2734, h-index is 29 and i-10 index 68 as of 14.01.2014.

He has participated in more than ten national and international research and innovations programs and has served as reviewer for many international journals, conferences, publishing companies and research funding organizations. Currently he coordinates the strategic the research project CoSound (2012–2015) and participate in three other research projects.

As regards synergistic activities, he took part in conference organizations, among other the IEEE Workshop on Machine Learning for Signal Processing (formerly Neural Networks for Signal Processing) 1999–2012. He is director of the Danish Sound Innovation Network (2009–2014), past chair of the IEEE Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2005–2007), and chair of IEEE Denmark Section's Signal Processing Chapter (2002–). He is a senior member of The Institute of Electrical and Electronics Engineers.



Tieniu Tan received his B.Sc. degree in electronic engineering from Xi'an Jiaotong University, China, in 1984, and his MSc and PhD degrees in electronic engineering from Imperial College London, U.K., in 1986 and 1989, respectively. He is a professor at National Laboratory of Pattern Recognition (NLPR), Institute of Automation of the Chinese Academy of Sciences (CAS). He also serves as Deputy Secretary-General of the CAS. His current research inter-

ests include biometrics, image and video understanding, and information forensics and security. He is a Member of the Chinese Academy of Sciences, and a Fellow of the IEEE and the IAPR (the International Association of Pattern Recognition). He currently serves as President of the IEEE Biometrics Council, First Vice President of the IAPR and Deputy President of the Chinese Association for Artificial Intelligence.