Special Issue on Subspace and Manifold Learning for Image and Video Indexing and Search

WITH THE rapid expansion of video and image data repositories available on the Internet, desktops and mobile devices, and multimedia indexing and searching technologies are becoming more and more important and have gained much attention from both academia and industry. Thousands of commercial TV channels and millions of individual grassroots producers, as well as billions of people, are sharing and consuming video and image content. The sheer size and complexity of multimedia repositories and the number of people creating, sharing, and consuming image and video content present a challenge to applications like image/video search (by either textual queries or image queries), browsing, content protection, duplication detection, commercial-ad insertion, and content tracking.

A key to enabling these applications is to find effective and compact representations that capture the essence of visual, syntactic, and semantic information conveyed in the image and video and have an efficient and scalable indexing scheme that supports fast search and mining operations. Subspace and manifold learning has gained significant advances lately for this purpose, particularly in graph embedding, vector subspace modeling, tensor modeling, canonical correlation analysis, etc.

In addition, the dimensionality of these subspace models is typically high; thus, sufficient information can be well preserved. However, high dimensionality brings us difficulties in efficiently indexing and searching data. How to create a compromise between effectiveness and robustness of the feature representation, with the underlying data organization and searching solution, is also a challenging issue. Approximate solutions, such as locality sensitive hashing and approximate search schemes that offer tradeoffs between accuracy and response speed, are promising directions to tackle this difficulty.

Moreover, image and video content does not exist in isolation, particularly for professionally produced media content. For example, a rich set of side information such as closed captions, audio track, and associated program guide, as well as usage statistics, provides crucial information for video search and mining. Due to the popularity of social network applications, online images and video clips are often associated with user-input tags, viewers' comments, as well as other textual information on the Web pages that host the images. How to develop cross-modal modeling and fusion solutions that facilitate more efficient and effective search, bridging the semantic gap, via subspace modeling and learning, is another interesting set of problems.

This Special Issue aims at the emerging applications of image and video indexing and search. It particularly focuses on the novel design and methodology of subspace and manifold learning, which are the best ways to foster original research. In total, 15 original contributions were submitted and evaluated following a rigorous review process. Eventually, four papers were accepted, resulting in an acceptance rate of 26.7%.

Wang *et al.* proposed a semisupervised classification algorithm using manifold regularization. A regularization trick is used to balance a tradeoff between loss and penalty in the objective function. This algorithm can fit the entire path of solutions for every value of the regularization hyperparameters with superior computational efficiency. To overcome the challenges of content variations and spatial dependence in visual data mining, Yuan *et al.* proposed to discover visual collocation patterns based on frequent itemset mining. A self-supervised subspace learning method is presented to refine the visual codebook by feeding back discovered patterns, which can effectively detect semantically meaningful patterns. Tian *et al.* proposed a unified multilinear framework to generalize TensorFace, i.e., view-manifold-based TensorFace and kernelized TensorFace, so that the local distances in the multiview face space and structure of the latent view manifold are all preserved to capture the nonlinearity of the view subspace. It advances the multiview face recognition performance in standard evaluations. Bian *et al.* proposed a transfer topic model for human action recognition from video in the scenario of insufficient training data. The basic idea is to use extra information obtained from an auxiliary domain to assist video representation in the target domain. As each human action is modeled as a mixture of topics, it defines Kullback–Leibler-divergence-based regularization by learning topics from the auxiliary domain to tune the topic estimation in the target domain, which improves the generalization of the cross-domain model.

We would like to thank all authors for their contributions and all reviewers for their timely service. We greatly acknowledge the strong support from Editor-in-Chief Prof. Eugene Santos and all kinds of help from his editorial colleagues.

YUN FU, *Guest Editor*
University at Buffalo, The State University of New York
Buffalo, NY 14260 USA

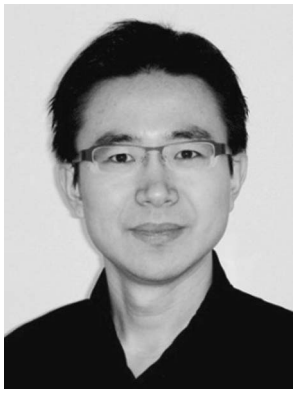
XIAN-SHENG HUA, *Guest Editor*
Microsoft Research
Seattle, WA USA

ZHU LI, *Guest Editor*
Futurewei (Huawei) Technologies
Bridgewater, NJ USA

AGGELOS K. KATSAGGELOS, *Guest Editor*
Northwestern University
Evanston, IL 60208 USA

THOMAS S. HUANG, *Guest Editor*
University of Illinois at Urbana–Champaign
Urbana, IL 61801 USA

Digital Object Identifier 10.1109/TSMCB.2011.2179292



Yun Fu (S'07–M'08–SM'11) received the B.Eng. degree in information engineering and the M.Eng. degree in pattern recognition and intelligence systems from Xi'an Jiaotong University, Xi'an, China, in 2001 and 2004, respectively, and the M.S. degree in statistics and the Ph.D degree in electrical and computer engineering from the University of Illinois, Urbana, in 2007 and 2008, respectively.

He was a research intern with Mitsubishi Electric Research Laboratories, Cambridge, MA, in summer 2005 and with the Multimedia Research Laboratory, Motorola Laboratories, Schaumburg, IL, in summer 2006. In 2008, he joined BBN Technologies, Cambridge, as a Scientist. In the spring of 2009, he was a part-time Lecturer with the Department of Computer Science, Tufts University, Medford, MA. Since 2010, he has been an Assistant Professor with the Department of Computer Science and Engineering, University at Buffalo, The State University of New York, Buffalo. His research interests include interdisciplinary research in machine learning, social media analytics, human–computer interaction, and cyber-physical systems.

Dr. Fu is a Life Member of the Association for Computing Machinery, Society of Photo-Optical Instrumentation Engineers, Association for the Advancement of Artificial Intelligence, and Institute of Mathematical Statistics and a Beckman Graduate Fellow in 2007–2008. He is the Associate Editor for IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He was the recipient of the 2002 Rockwell Automation Master of Science Award, Edison Cups of the 2002 GE Fund Edison Cup Technology Innovation Competition, the 2003 Hewlett-Packard Silver Medal and Science Scholarship, the 2007 Chinese Government Award for Outstanding Self-financed Students Abroad, the 2007 IEEE International Conference on Image Processing Best Paper Award, the 2007-2008 Beckman Graduate Fellowship, the 2008 M. E. Van Valkenburg Graduate Research Award, the ITESOFT Best Paper Award of the 2010 International Association for Pattern Recognition International Conferences on the Frontiers of Handwriting Recognition, the 2010 Google Faculty Research Award, the 2011 IEEE International Conference of Multimedia and Expo Quality Reviewer, the Large Scale Visual Analytics (LSVA) Best Paper Award of the 2011 IEEE International Conference on Data Management Workshop on LSVA, and the 2011 IC Postdoctoral Research Fellowship Award.

Xian-Sheng Hua (M'05) received the B.S. degree and the Ph.D. degree in applied mathematics from Peking University, Beijing, China, in 1996 and 2001, respectively.

In 2001, he joined Microsoft Research Asia as a Researcher. Since then, his research interests have been in the areas of multimedia search, advertising, understanding, and mining, as well as pattern recognition and machine learning. He has authored or coauthored more than 200 research papers in these areas and has filed more than 60 patents. He is currently with Microsoft Research, Seattle, WA, where he became the Principal Research and Development Lead in Multimedia Search for the Microsoft search engine, Bing, in 2011. He oversees a team that designs and delivers leading-edge media understanding and indexing features. He is currently an Adjunct Professor with the University of Science and Technology of China, Hefei, China. He serves as an Associate Editor of the *Association for Computing Machinery (ACM) Transactions on Intelligent Systems and Technology*, an Editorial Board Member of *Advances in Multimedia* and *Multimedia Tools and Applications*, and an Editor of *Scholarpedia* (multimedia category).

Dr. Hua is a member of the Video Signal Processing and Communications Technical Committee (TC) and Multimedia Systems & Applications (MAS) TC of IEEE Circuits and Systems Society, the Chair of the Interest Group on Visual Analysis and Content Management in Multimedia Communication TC of IEEE Communications Society, and a senior member of ACM. He serves as an Associate Editor of IEEE TRANSACTIONS ON MULTIMEDIA. He was the Vice Program Chair of the 2005 Visual Communications and Image Processing; the workshop organizer of the 2009 and 2010 IEEE International Conference of Multimedia and Expo (ICME), 2009 IEEE International Conference on Data Management, and 2010 ACM Multimedia workshops; a senior member of the Technical Program Committee and the Area Chair of ACM Multimedia and ACM Knowledge Discovery and Data Mining (KDD); and the Demonstration Chair, the Tutorial Chair, the Special Session Chair, and a program committee member of many more international conferences. He will serve as a Program Cochair for the 2012 ACM Multimedia and 2012 IEEE ICME. He was honored as one of the recipients of the prestigious 2008 MIT Technology Review TR35 Young Innovator Award for his outstanding contributions to video search. He won the Best Paper and Best Demonstration Awards at the 2007 ACM Multimedia, the Best Poster Award at the 2008 IEEE International Workshop on Multimedia Signal Processing, the Best Student Paper Award at the 2009 ACM Conference on Information and Knowledge Management, and the Best Paper Award at the 2010 International Conference on MultiMedia Modeling. He was named one of Global Entrepreneur's "Business Elites of People under 40 to Watch" in 2009. He received six Microsoft "Ship-It" awards for inventing and shipping six technologies into Microsoft mainstream products.

Zhu Li (SM'07) received the Ph.D. degree in electrical and computer engineering from Northwestern University, Evanston, IL, in 2004.

He was an Assistant Professor with the Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong, from 2008 to 2010 and a Senior Research Engineer, Senior Staff Research Engineering, and then Principal Staff Research Engineer with the Multimedia Research Laboratory, Motorola Laboratories, Schaumburg, IL, from 2000 to 2008. He is currently a Senior Staff Researcher with the Media Networking Laboratory, Core Networks Research, Futurewei (Huawei) Technologies, Bridgewater, NJ, where he leads the Media Analytics and Processing Group. He is a Coeditor for the Springer-Verlag book on "Intelligent Video Communication: Techniques and Applications." His research interests include audiovisual analytics and machine learning with its application in large-scale video repository annotation, and search and recommendation, as well as video adaptation, source-channel coding, and distributed optimization issues of wireless video networks. He has 20 issued or pending patents and more than 70 publications in book chapters, journals, conference proceedings, and standards contributions in these areas.

Dr. Li is an elected Vice Chair of the IEEE Multimedia Communication Technical Committee in 2008-2010. He served on numerous conference and workshop Technical Program Committees and on the Best Paper Award Committee for the 2010 IEEE International Conference of Multimedia and Expo (ICME) and was a symposium Cochair of the 2008 IEEE International Conference on Communications. He was the recipient of the Best Poster Paper Award from IEEE ICME in Toronto, ON, Canada, in 2006 and the Best Paper Award from IEEE International Conference on Image Processing in San Antonio, TX, in 2007.

Aggelos K. Katsaggelos (S'80–M'85–SM'92–F'98) received the Diploma degree in electrical and mechanical engineering from the Aristotelian University, Thessaloniki, Greece, in 1979 and the M.S. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1981 and 1985, respectively.

In 1985, he joined the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, where he is currently a Professor holder of the AT&T chair. He was previously the holder of the Ameritech Chair of Information Technology (in 1997–2003). He is also the Director of the Motorola Center for Seamless Communications, a member of the Academic Staff of the NorthShore University Health System, and an affiliated Faculty Member with the Department of Linguistics, Northwestern University, and he has an appointment with the Argonne National Laboratory. He has published extensively in the areas of multimedia signal processing and communications (over 180 journal papers, 400 conference papers, and 40 book chapters), and he is the holder of 19 international patents. He is the coauthor of *Rate-Distortion Based Video Compression* (Kluwer, 1997), *Super-Resolution for Images and Video* (Claypool, 2007), and *Joint Source-Channel Video Transmission* (Claypool, 2007).

Prof. Katsaggelos is a fellow of the Society of Photo-Optical Instrumentation Engineers (in 2009). Among his many professional activities, he was the Editor-in-Chief of the IEEE SIGNAL PROCESSING MAGAZINE (in 1997–2002), a Board of Governors member of the IEEE Signal Processing Society (in 1999–2001), and a member of the Publication Board of the IEEE Proceedings (in 2003–2007). He was the recipient of the IEEE Third Millennium Medal (in 2000), the IEEE Signal Processing Society Meritorious Service Award (in 2001), the IEEE Signal Processing Society Technical Achievement Award (in 2010), an IEEE Signal Processing Society Best Paper Award (in 2001), an IEEE International Conference of Multimedia and Expo Paper Award (in 2006), an IEEE International Conference on Image Processing Paper Award (in 2007), and an International Symposium on Image and Signal Processing and Analysis Paper Award (in 2009). He was a Distinguished Lecturer of the IEEE Signal Processing Society (in 2007-2008).

Thomas S. Huang (S'61–M'63–SM'76–F'79–LF'01) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, and the M.S. and Sc.D. degrees in electrical engineering from Massachusetts Institute of Technology (MIT), Cambridge.

He was a Faculty Member with the Department of Electrical Engineering, MIT, from 1963 to 1973 and a Faculty Member with the School of Electrical Engineering and the Director of its Laboratory for Information and Signal Processing, Purdue University, West Lafayette, IN, from 1973 to 1980. In 1980, he joined the University of Illinois, Urbana, where he is currently the William L. Everitt Distinguished Professor of Electrical and Computer Engineering, a Research Professor with the Coordinated Science Laboratory, the Head of the Image Formation and Processing Group, Beckman Institute for Advanced Science and Technology, and the Cochair of the institute's major research theme: human–computer intelligent interaction. He is a Founding Editor of the *International Journal of Computer Vision, Graphics, and Image Processing* and the Editor of the *Springer Series in Information Sciences* (Springer). He has published 20 books and over 500 papers in network theory, digital filtering, image processing, and computer vision. His professional interest includes the broad area of information technology, particularly the transmission and processing of multidimensional signals.

Dr. Huang is a member of the National Academy of Engineering, a foreign member of the Chinese Academies of Engineering and Science, and a fellow of the International Association of Pattern Recognition and the Optical Society of America. He has received a Guggenheim Fellowship and a fellowship from the Japan Association for the Promotion of Science. He was the recipient of an A. von Humboldt Foundation Senior U.S. Scientist Award, the IEEE Signal Processing Society's Technical Achievement Award in 1987, the Society Award in 1991, the IEEE Third Millennium Medal in 2000, the Honda Lifetime Achievement Award for his "contributions to motion analysis" in 2000, the IEEE Jack S. Kilby Medal in 2001, the King-Sun Fu Prize from the International Association of Pattern Recognition in 2002, and the Pan Wen-Yuan Outstanding Research Award.