Visual Communications and Image Processing

Cheng-Tie Chen
Bellcore, NVC 3X-321
331 Newman Springs Road
Red Bank, New Jersey 07701
E-mail: ctc@nyquist.bellcore.com

Kou-Hu Tzou
COMSAT Laboratories
22300 Comsat Drive
Clarksburg, Maryland 20871
E-mail: kouhu@ctd.comsat.com

Ya-Qin Zhang
David Sarnoff Research Center
201 Washington Road
Princeton, New Jersey 08543-5300
E-mail: zhang@earth.sarnoff.com

Video and multimedia communication is a natural evolution from conventional voice communication. The last decade has witnessed phenomenal progress in the field in terms of research and development, products, services, and standardization activities. Advances in new image processing algorithms, video/multimedia communication standards, computer technologies VLSI devices, and fiber optics technologies will definitely continue to grow rapidly in association with global information superhighway construction and will create many valuable video services to the vast number of customers. Therefore, this fifth special section on visual communications and image processing is published to update the progress in this field.

A total of 18 papers has been collected for this special section. Video/image coding still shows a strong interest with eight papers in this category covering standard activities; automata; vector quantization; and subband, wavelet, and object-oriented coding. Five papers deal with motion estimation issues. The remaining papers deal with image processing algorithms and applications.

The two invited papers provide excellent overviews of the two important standards activities in audio-visual coded representation undertaken by ISO and ITU. Reader, the chairman of the ISO/MPEG4, discusses and outlines the background, vision, applications, scope of work, and workplan of MPEG4, a continuing effort of the ISO/WG11 after its successful completion of MPEG1 and MPEG2. The paper by Schaphorst, the rapporteur for very low bitrate visual telephony for ITU-T/SG15, describes the recent effort in H.324 to produce a video telephony standard in the Public Switched Telephone Network (PSTN) and the mobile environment.

Quenneville and Meunier draw a parallel between quadtree and rational languages that are recognizable by finite state automata in their paper. The authors illustrate that the automata construction by recursive decomposition of the image is similar to the construction of region-based quadtree. Examples show how to apply the method to binary images, gray-scale images, color images, and 3-D images.

The paper by Martucci received the Best Young Investigator Award paper of the 1994 Visual Communication and Image Processing conference. The author shows that using discrete sine or cosine transforms (DST or DCT) to implement a filter on an image is equivalent to conventional linear filtering except at the image boundaries, where a symmetric or asymmetric extension is implied. Therefore, DSTs and DCTs become an alternative to the discrete Fourier transform for transform domain filtering with the advantages of real number arithmetic and abundant software/hardware available.

The paper by Mandal et al. proposes several fast algorithms for wavelet-domain motion estimation, which have improved the coding efficiency over the baseline multiresolution motion estimation. Tzovaras, Grammalidis, and Strintzis present an object-based method for 3-D video coding, using motion and disparity information between the left and right channel. Xie, Eycken, and Oosterlinck describe a new technique for hierarchical motion estimation with smoothness constraints. The paper by Kim and Park proposes a new mapping parameter technique for motion estimation that can be used for objected-oriented analysis-synthesis video coding. It is reported that its performance is better than the conventional gradient approach. Temporal rate up-conversion is a scalability tool in MPEG2. The paper by Huang and Chao proposes a new scan-rate up-conversion method based on frame-based motion-compensated interpolation for moving objects. Better performance is reported by using this method.

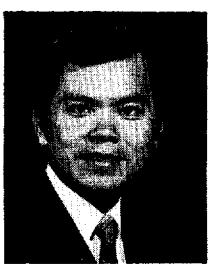
Segmented and object-oriented image coding is an important MPEG4 activity. The paper by Philips introduces a new class of weakly separable orthonormal bases for object segmentation. The author shows that the proposed method greatly reduces computational complexity at the expense of slight

degradation, as compared to the conventional Gram-Schmidt procedure.

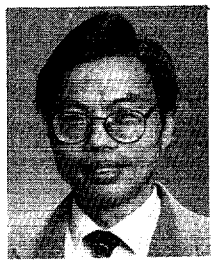
Predictive residual vector quantization has been developed recently to overcome the complexity problem of vector quantization (VQ). Rizvi, Nasrabadi, and Wang present a new scheme that imposes a constraint on the output bit rate. Specifically, the authors designed two-, four-, and six-stage entropy-constraint predictive residue VQ with the fixed bit rates of 0.75, 1.5, and 2.25 bpp, respectively. Simulation results based on several images showed that the proposed method may achieve up to a 50% bit rate savings over the JPEG algorithm. The paper by Hu, Wang, and Wang applies feature extraction techniques to photo ID compression. A region of facial features, containing the eyes and mouth, is detected and renders more accurately than the rest of the image. The technique can compress a 128×128×8 bits photo ID image to an average of 350 bytes.

Kim et al. present a wavelet image coding scheme using a tree structure that exploits the interband geometric correlation with the incorporation of human visual characteristics. Ma and Rajala introduce a subband-based block truncation coding algorithm with various bit allocation methods. Zhou, Chang, and Han discuss the admission control and bandwidth allocation issue for multimedia communication in ATM networks.

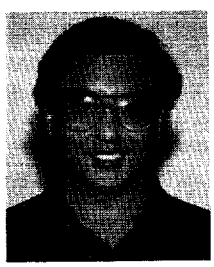
Tom, Lay, and Katsaggelos apply the expectation-maximization algorithm to multichannel simultaneous blur identification and restoration of images. The paper by Blohm deals with a fundamental problem of scene illumination and display (presentation) in video signal processing. The method compresses the dynamic range of physical light intensity into a finite number of bits that give better visual perception for portrait images. Measuring a 3-D object size using machine vision technique has been a practical application in the field. The paper by Hsu and Fuh develops a method to improve performance by using local and global thresholding, K-cosine algorithm for edge terminal extraction, and terminal extension to fill gaps.



Cheng-Tie Chen received the BSEE degree from National Taiwan University in 1977, and MSE and PhD degrees from the Moore School of Electrical Engineering, University of Pennsylvania, in 1981 and 1983, respectively. From 1983 to 1984, he was employed as a research associate at the Department of Electrical Engineering, Princeton University, where he was engaged in the research of non-Gaussian and sensor-array signal processing. From 1984 to 1987, he was working for Research Laboratories, Eastman Kodak Company as a senior research scientist. His research activities there were in the areas of image restoration and image coding/transmission. Since 1987, he has been with Bell Communications Research in New Jersey, as a member of the technical staff. His current research interests are in the areas of visual communication networking and multimedia applications.



Kou-Hu Tzou is currently a senior scientist at COMSAT Laboratories. He received his BS degree in telecommunication engineering from the National Chiao-Tung University, Taiwan, in 1975, and MS and DSc degrees in electrical engineering from Washington University, Saint Louis, Missouri, in 1980 and 1983, respectively. Since 1991, he has been with COMSAT Laboratories in Clarksburg, Maryland, where he is responsible for research and development in the areas of television signal processing, digital video, and HDTV transmission via satellite. He was with Bellcore in Red Bank, New Jersey, from 1987 to 1991 as a member of the technical staff in the Video Signal Processing Research Group, where he engaged in HDTV coding for transmission over the fiber-based broadband ISDN. Prior to joining Bellcore, Tzou was with GTE Laboratories in Waltham, Massachusetts, where he was a principal investigator of the Image Processing Group. Tzou has published more than 50 refereed conference and journal papers in the areas of digital communications, image compression, video coding and transmission, and hardware architecture for video signal processing. He is a corecipient of the 1993 IEEE Video Technology Best Paper Award. He taught several short courses on digital image coding and transmission and HDTV signal processing. From 1989 to 1991, he served as associate editor for *IEEE Transactions on Circuits and Systems*. Currently, he serves as associate editor for *IEEE Transactions on CAS for Video Technology*. He also served as a co-guest editor for the *Optical Engineering* special sections on visual communications and image processing over the past few years. He was the conference chair of the 1991 SPIE Visual Communication and Image Processing conference in Boston, Massachusetts. He taught a short course on digital video at University of California, Berkeley Extension Program in 1994 and 1995. Tzou is an IEEE Senior Member and an SPIE Fellow.



Ya-Qin Zhang received his BS and MS in electrical engineering from the University of Science and Technology of China (USTC) in 1983 and 1985, respectively. He received his PhD in electrical engineering from George Washington University, Washington D.C., in 1989. He is the head of digital video communications at David Sarnoff Research Center in Princeton, New Jesey. His group is actively involved in the research and development of video compression algorithms and products for Grand Alliance HDTV, DirecTV, MPEG2 codec, and video-telephony applications. He wasresponsible for the video compression activities in the Applications Technologies and Systems Department at GTE Laboratories, Inc., Waltham, Massachusetts. He has authored and coauthored more than 100 refereed papers (including 30 in various IEEE Transactions journals) and several U.S. patents in image/video compression and communications, wireless networking, satellite communications, and medical imaging. He was an adjunct faculty member at George Washington University and Tufts University in 1990 and 1994. Zhang serves as the Editor for Express Letters for the *IEEE Transactions on Circuits and Systems for Video Technology*. He was a co-guest editor for the special issue on Advances in Image and Video Compression for the Proceedings of IEEE in February 1995. He serves on the editorial boards of seven other professional journals and more than a dozen conference committees. He currently represents Sarnoff (and GTE between 1990 to 1994) in the ISO/MPEG and ITU-T/LBC standardization efforts. He is a Senior Member of IEEE and a member of Eta Kappa Nu.