Visual Communications and Image Processing IV

Cheng-Tie Chen
Bellcore
445 South Street
Morristown, New Jersey 07962

Hsueh-Ming Hang
National Chiao-Tung University
Center for Telecommunication Research
Hsinchu, Taiwan

Kou-Hu Tzou, FELLOW SPIE
COMSAT Laboratories
22300 Comsat Drive
Clarksburg, Maryland 20871

Over the past few years the worldwide R&D activities in visual communications and image processing have continued to grow vigorously. Therefore, we are pleased to present this fourth special section on this topic in *Optical Engineering*. We accepted 23 papers from various areas including wavelet/subband processing, motion estimation, vector quantization, multiresolution processing, DPCM, progressive transmission, neural networks, advanced video processing, pattern recognition, and applications of digital image processing.

The first paper, by Haddad, Akansu, and Benyassine, presents a tutorial on time-frequency localization in transforms, subbands, and wavelets. This paper reviews recent extensions of the "uncertain" principle to the discrete case and develops optimum wave forms. It also reviews common features of block transforms, subband filter banks, and wavelets and demonstrates how the discrete uncertainty can be used to evaluate these decomposition methods. The paper by Balakrishnan and Pearlman presents a hexagonal subband image coding scheme in which a hexagonally sampled image is decomposed into a lowpass band and nine other bands of one octave bandwidth and 60-deg angular orientation. The rate is allocated based on a distortion measure weighted by the human visual system frequency response. Interesting results, both theoretical and experimental, can be found in this paper. The third paper in this category is presented by Nicoulin et al. It contains two parts—a filter bank design for subband analysis/synthesis and an image compression scheme composed of a scalar quantizer with vector-quantization-based entropy coding. New ideas and good results are demonstrated.

In the area of motion analysis/estimation, we accepted four papers. The paper by Young and Kingsbury introduces an interesting motion estimation technique based on the lapped transform. The unique feature is that the transform is applied for both motion estimation and coding. Seferidis and Ghanbari develop a general approach to motion estimation that can cope with rotation as well as other nonlinear deformations. The paper by Blanc-Féraud, Barlaud, and Gaidon deals with discontinuities of the motion field that often occur at the edges of a moving object. The authors derive a new deterministic relaxation algorithm in a coarse-to-fine resolution scheme and achieve more accurate results. A motion classification scheme using a 3-D subband structure is developed by Ngan and Chooi. It is able to classify the motion into eight velocity classes and several orientations within each class. This scheme has been tested on real-life video sequences, and subjective evaluations confirm the validity of the results.

Wu and Gersho describe in their paper a lapped decoding technique to improve vector quantization by exploiting the interblock correlation at the decoder. This technique leads to higher SNR and better perceptual quality as well. Wen and Lu propose an image compression scheme that combines classified vector quantization (VQ) in conjunction with discrete cosine transformation (DCT) and block truncation coding (BTC). Edge areas are segmented by considering four DCT coefficients at the lowest frequencies. The edge areas are coded with VQ/BTC and the smooth areas with VQ/DCT. Wen and Lu's results show that good quality is achieved at 0.4 bit/pixel for a variety of images. Mohsenian, Rizvi, and Nasrabadi present a predictive VQ scheme in which the current vector is predicted from the neighboring four vectors. The prediction is achieved based on a neural network approach. The VQ in association with a Kohonen codebook is applied to code the prediction error vector. They demonstrate high performance for the proposed predictive VQ scheme.

DPCM is a well-known technique widely used for speech and video coding because of the advantage of low system complexity. Daut, Zhao, and Wu develop a new double predictor DPCM algorithm to improve coding efficiency and robustness. Hwang, Venkatraman, and Rao continue their effort to encode images with the human visual system (HVS) in the transform coding structure and present their recent results using the lapped orthogonal transform together with classified VQ. With the techniques discussed, the block artifacts that often appear in conventionally transform-coded images are significantly reduced. Additionally, a progressive transmission technique is included to facilitate interactive communications.

In the paper by Chang, Lee, and Horng, a Hopfield neural network approach for finding trajectories of feature points in a monocular image sequence is described. Using the massive parallel-processing power of neural networks, an accurate solution can be obtained in real time. Another neural network paper is presented by Ansari and Liu. They describe a high-speed partial shape recognition system based on a bidirectional associative memory (BAM) neural network. Objects are represented by landmarks, whose values are trained and stored in the neural network. A fast and rather accurate recognition of partially occluded objects can then be performed by this network.

In the paper by McLaren and Gersho, a multiresolution HDTV compression algorithm for video tape recording is proposed. The scheme allows for the variable speed playback of recorded material at speeds of up to 100 times the normal speed in both the forward and reverse directions.

Dufaux et al. propose a method for generic motion-compensated video coding for both progressive and interlaced input. They consider a quadtree-based multigrid motion estimation to adapt the local motion more faithfully. The motion-compensated difference is then decomposed into a multiresolution data structure via a Gabor-like wavelet transform, followed by an arithmetic coder. They show that good quality is achieved for a wide range of bit rates by the use of this generic coding scheme.

Model-based coding is considered to be a candidate for the future MPEG4 standard. Huang, Ouhyoung, and Wu present an automatic feature point extraction scheme to establish the wireframe model used in model-based coding. Good results are obtained with the application of this automatic extraction scheme to face images. A 2-D segmentation-based coding scheme is proposed by Kwon and Chellappa. Images are partitioned into uniform and textured regions. Then, the uniform regions are represented by polynomials and the texture regions by a Gauss-Markov random field model. A high compression ratio has been achieved with acceptable picture quality. Only recently has a fractal model been applied to video sequences for compression purposes. A 3-D fractal-based video coding scheme is suggested by Li, Novak, and Forchheimer. They generalize the affine transform to a more general quadratic form and apply it to 3-D image blocks. The initial results reported indicate that this approach has good potential.
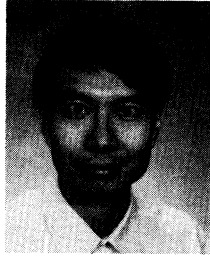
The paper by Hsu and Tsai presents an image compression scheme that detects edges based on moment preservation. The detection scheme is applied to image compression by the approximation of the detected edges into piecewise linear segments. They report good performance when the scheme is applied to images of Chinese characters and trademarks. In the paper by Sun et al., a color image analysis technique is applied to liver tissue classification. The system coarsely segments the tissue image using the achromatic information and then makes use of the chromatic information to classify the segmented region into different tissue classes. Results based on different sample sets show this technique to be reliable.

Pei and Hsu present color pattern recognition with the use of complex log mapping. First, 3-D color images are projected onto a 2-D color plane, which is followed by a log mapping transform for recognition. They show that the method can effectively recognize color objects subject to scaling and rotation varieties. Character recognition with high accuracy is a difficult task, especially when the number of possible characters is large. Miou et al. implement a Chinese character recognition system with a noble algorithm that includes feature extraction unique to Chinese characters. To speed up the recognition, they designed a pipelined VLSI for the most time-consuming matching and sorting processes. They report that with the use of the VLSI chip, the system can recognize 200 characters/s, which is 100 times faster than the software-based approach.

We would like to thank the authors who contributed to this special section for their efforts and cooperation, which have made this publication possible. We also are grateful to the many reviewers for their diligence and kind help.
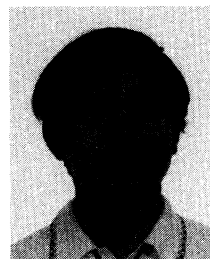
Cheng-Tie Chen received the BSEE degree from National Taiwan University in 1977 and the MSE and PhD degrees from the Moore School of Electrical Engineering, University of Pennsylvania, in 1981 and 1983, respectively. From 1983 to 1984, he was employed as a research associate at the Department of Electrical Engineering, Princeton University, where he was engaged in the research of non-Gaussian and sensor-array signal processing. From 1984 to 1987, he worked in the research laboratories at 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 commu-

Downloaded from SPIE Digital Library on 19 Oct 2010 to 66.165.46.178. Terms of Use: http://spiedl.org/terms

Hsueh-Ming Hang received his BS and MS degrees from National Chiao-Tung University, Hsinchu, Taiwan, in 1978 and 1980, respectively, and his PhD in electrical engineering from Rensselaer Polytechnic Institute, Troy, New York, in 1984. From 1984 to 1991, he worked at AT&T Bell Laboratories, Holmdel, New Jersey, where he was engaged in digital image compression algorithm and architecture research. He joined the Electronics Engineering Department of National Chiao-Tung University, Hsinchu, Taiwan, in December 1991. His current research interests include digital video compression, image/signal processing algorithms and architectures, and digital communications theory. Dr. Hang holds three patents and has published over 30 technical papers related to image compression, signal processing, and video codec architecture. He has served as a session chair or a conference co-chair for several international conferences on visual communications and image processing. He is currently an associate editor of *IEEE Transactions on Image Processing* and is a senior member of IEEE and a member of Sigma Xi.
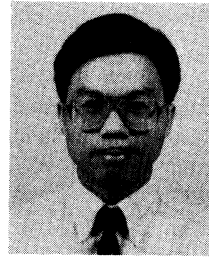
Kou-Hu Tzou received his BS degree in telecommunications engineering from National Chiao-Tung University, Hsinchu, Taiwan, in 1975, and his MS and DSc degrees in electrical engineering from Washington University, St. Louis, Missouri, in 1980 and 1983, respectively. Since October 1991, he has been with COMSAT Laboratories in Clarksburg, Maryland, as the manager of the Image Processing Department, 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, 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 Bellcore, Dr. Tzou was with GTE Laboratories in Waltham, Massachusetts, where he was a principal investigator of the Image Processing Project. Dr. Tzou has published over 50 conferences and journal papers in the areas of digital communications, image compression, video coding and transmission, and hardware architecture for video signal processing. He has served as an associate editor on *IEEE Transactions on Circuits and Systems,* and currently serves as an associate editor on *IEEE Transactions on Circuits and Systems for Video Technology.* He is the current chair-elect of the Visual Signal Processing and Communication Technical Committee, IEEE Circuits and Systems Society. Dr. Tzou is an IEEE Senior Member and an SPIE Fellow.