

Special Section on Image/Video Compression and Processing for Visual Communications

Yao Wang

Polytechnic University
Department of Electrical Engineering
Brooklyn, New York 11201
E-mail: yao@vision.poly.edu

Touradj Ebrahimi

Swiss Federal Institute of Technology-EPFL
Signal Processing Laboratory
CH-1015 Lausanne, Switzerland
E-mail: Touradj.Ebrahimi@epfl.ch

Majid Rabbani

Eastman Kodak Company
1700 Dewey Avenue
Rochester, New York 14650-1816
E-mail: rabbani@kodak.com

Ya-Qin Zhang

Sarnoff Corporation
Multimedia Technology Laboratory
201 Washington Road, CN5300
Princeton, New Jersey 08543-5300
E-mail: yzhang@sarnoff.com

This special section aims to expose to the readers recent advances in image and video compression and processing technologies for visual communications. Given the very broad scope of the field, this special section focuses on the following areas:

1. New theory and algorithms for image and video compression, with focus on compression for image and video databases, very high-quality and lossless image compression, compression of mixed documents, pre- and post-processing in lossy compression and communication systems;
2. emerging areas in digital visual communication, including digital watermarking and content-based processing;
3. system level descriptions of applications utilizing above techniques, such as medical imaging,

digital banking, digital museum and library, surveillance and access control, image and video distribution over Internet and other networks; and

4. review of the status of recent and emerging standards.

A total of 20 papers, selected by the peer-review process, are presented here in three parts. Part I includes nine papers presenting new progress in image compression, with focus on algorithms that enable progressive transmission and content-based retrieval.

The first paper, "Decoding compression with reversible embedded wavelets (CREW) codestreams," by Boliek et al., reviews the CREW algorithm which is a proposal submitted in response to the JPEG2000 Standard Committee's recent call for proposals. A detailed description of the features

and functionalities of the CREW compression algorithm is provided here.

The second paper, "High quality document image compression with 'DjVu,'" by Haffner et al., describes a new scheme for compression of high resolution color document images. The proposed scheme separates the text/graphics (foreground) in a document image from the texture background. The foreground mask is compressed at high spatial resolution, using a new bi-level image coding technique. The background image is coded at a lower resolution, using a new wavelet-based method that does not waste bits to specify coefficients that correspond to foreground pixels. The developed method can achieve very high compression, reducing a typical magazine page from 23 MB to about 50 KB while retaining very high perceptual quality.

In addition, the compression format allows progressive retrieval of a document.

In the next paper, "Check image compression using a layered coding method," Huang, Wang, and Wong consider the compression of a special type of document images—check images. Similar to the previous paper, the authors propose to decompose a check image into several layers of foreground and background information and code each layer according to its characteristics. The proposed technique is shown to produce images of better quality than the traditional JPEG or wavelet coding methods applied to the entire image.

Swanson and Tewfik, in their paper "Fast progressively refined image retrieval," describe a new image compression scheme that facilitates fast query matching and progressive retrieval. The scheme stores the "keywords" for an image in the beginning of the compressed file as a header. The locations and sizes of these "keywords" in the image is stored next. Finally, the non-object regions are coded using the JPEG algorithm. The wavelet images corresponding to keywords are reduced to a finite set using vector quantization. The codewords and the ordering for the keywords are designed to achieve a desirable tradeoff between the file size and the query response time. The approach enables fast image retrieval and browsing, while maintaining a high compression efficiency (5–10% lower than the JPEG standard).

Ramos and Hemami, in "Perceptually based scalable image coding for packet-based networks," propose an image coding scheme that offers perceptually driven scalability. Scalability and perceptual coding are combined through intraband coding of perceptually significant image regions. Their approach is especially suitable for applications requiring high quality images at several spatial resolutions and bitrates in heterogeneous networks (databases, the Internet, etc.).

Moni, in "Application-specific image compression for multimedia applications," describes a set of wavelet-based compression schemes adapted to different types of images. Simulation shows promising results of the scheme.

Luo and Chen consider, in "Coherently three-dimensional wavelet-based approach to volumetric image com-

pression," the compression of three-dimensional images. They propose a coherently 3D approach to volumetric image compression by integrating a 3D wavelet transform, a pseudo-uniform quantization, and 3D zerotree coding. They apply this compression system to volumetric medical data (CT) and show significant improvement when compared to a 2D approach.

Ansari and Memon present, in "Near-lossless image compression techniques," compression techniques based on limiting the maximum allowable deviation of pixel values. Several different approaches are investigated and compared including modified lossless predictive coding techniques, reversible transforms in conjunction with pre-quantization, and a partially embedded two-layer scheme using a lossy wavelet transform followed by near-lossless coding of the residual.

The final paper in this group deals with the implementation issue in compression. In "Multi-DSP architecture for real time lattice quantization indexing," Moureaux, Nus, and Antonini consider the design of a multi-DSP architecture for a recently published lattice vector quantization (LVQ) indexing algorithm. The designed system allows an efficient trade-off between storage cost and complexity, thus providing an adaptive coding strategy well suited to many diverse applications.

Part II consists of seven papers dealing with video compression and the associated motion estimation problem.

The first paper, "MPEG-4 video verification model: a solution for interactive multimedia applications," by Fleury et al. provides a high-level overview of the MPEG-4 video verification model, which is the reference implementation of the MPEG-4 video encoder for iterative improvement of different core experiments.

The next paper, "On the coding of the enhancement layer of an MPEG-2 SNR scalable order," by Wilson, Pek, and Ghanbari, focuses on the SNR scalable coding mode of MPEG-2 standard, and presents solutions to improve the coding efficiency of the enhancement layer. Recognizing that the distribution of the enhancement layer coefficients is very different from that of a single layer and that it is dependent on whether the corresponding base layer coefficients are zero, the authors investigated the use of multiple arith-

metic and run-length coders. The best option can deliver up to 20–25% better compression in the enhancement layer than using the same Huffman coding table for the run-lengths of both base and enhancement layers.

The paper "Motion-compensated transform coding of video using adaptive displacement fields," by Bi et al. proposes a new scheme for optimizing the rate-distortion performance of a coder using a quad-tree motion representation. A novelty of the approach is that the motion compensation error corresponding to different block-sizes for representing the motion vectors are actually coded to determine their true impact on the overall rate-distortion cost. Improvements over the H.263 standard are demonstrated for low-bit video applications.

Wavelet-domain motion estimation is revisited in the paper "Video coding with a variable block-sizing technique in the wavelet transform domain," by Lee and Ngan, with new improvements in using variable block sizes obtained from quad-tree segmentation. Positive results were obtained at very low bit rates compared with the H.263 standard.

The paper "Two-layer video coding and cell-loss concealment using pyramid structures," by Hong, Park, and Lee describes a two-layer video coding scheme using a Laplacian–Gaussian pyramid. A cell-loss concealment method using the pyramid structure is also proposed. Several realizations are compared. Results show that including the top two levels (the top Gaussian image and the next Laplacian image) in a three level pyramid in the base-layer yields the best trade-off among the coding efficiency, the bustiness of the bit stream, and robustness to cell loss.

Wus and Li investigated the use of vector wavelet coding (VWC) for video compression in "Video coding using a vector wavelet framework." The efficiency of using a finite state VWC method is compared against that of motion-compensated prediction. In order to circumvent the problem of the training-coding loop associated with vector quantization, VWC is combined with lattice vector quantization. Promising results are shown.

Powerful but low-cost motion estimation is one of the key issues in the successful deployment of video coding, especially in the consumer market applications such as video-conferencing.

Oh and Lee, in "Block-matching algorithm based on dynamic adjustment of search window for low bit-rate video coding," present a novel motion estimation technique based on a dynamic search window which aims at reducing computational complexity of block matching, while maintaining performance close to a full search technique. Their technique compared favorably with other fast search algorithms in terms of the quality of obtained results. Being orthogonal compared to other available methods, this method can be easily combined with others to further reduce the complexity.

Part III includes four papers relating to image and video processing.

Yeung and Mintzer, in "Invisible watermarking for image verification," consider the problem of data security and protection in image-based communications and present a novel watermarking algorithm that hides invisible information within high quality color and grayscale images. They describe a mechanism to verify whether an image has been altered with respect to the hidden watermarks. They extensively discuss advantages of their proposed technique over other invisible watermarking methods, and provide several results for different applications. Recent proliferation of digital visual content has created a new problem of intelligent information management. Günsel, Ferman, and Tekalp, in "Temporal video segmentation using unsupervised clustering and semantic object tracking," propose a content-based temporal video segmentation system which integrates both syntactic (domain independent) and semantic (domain dependent) features. They present new techniques for scene change detection and shot classification. These methods have been successfully applied to classification of TV news units.

Murching and Woods consider pre- and post-processing for image coding in "Linear pre/post filters for DCT-based image coding systems." They develop a novel analytical technique for the design of spatially adaptive FIR pre/post filters for DCT-based image coding systems while using a quantitative objective function. Simulations demonstrate that such filters can help the DCT-based coders match the performance of subband/wavelet coders.

In their paper "On the motion compensation within a down-conversion decoder," Vetro and Sun consider the

decoding of an MPEG compressed high-resolution video for low-resolution display, which is required, for example, for receiving HDTV on a standard TV set. If one simply performs motion compensation based on the previously decoded low-resolution frames, the blurring caused by conventional bilinear interpolation will cause periodic impulsion in the decoded image known as drift. The authors propose to combine motion compensation with frequency-domain down-conversion and develop optimal motion-compensation filters which can minimize the mean square error of the down-conversion. Significant visual improvements are demonstrated.

We would like to take this opportunity to thank all the authors for contributing their latest research results to this special issue, and to the anonymous reviewers for their thorough and critical examination of the manuscripts. Finally, we would like to acknowledge the contribution of Professor Murat Kunt in the organization of this special section.



Yao Wang received her BS and MS degrees in electrical engineering from Tsinghua University, Beijing, China, in 1983 and 1985, respectively, and her PhD degree in electrical engineering from the University of California at Santa Barbara in 1990. Since 1990 she has been on the faculty of Polytechnic University, Brooklyn, New York, and is presently an associate professor of electrical engineering. From 1992 to 1996 she was a consultant with AT&T Bell Laboratories, Holmdel, New Jersey. Since 1997 she has continued to be a consultant with AT&T Labs Research, Red Bank, New Jersey. This year she is on sabbatical leave at Princeton University. Since 1995 she has been serving as Associate Editor for *IEEE Transactions on Circuits and Systems for Video Technology* and for the *Journal of Visual Communications and Image Representation*. She is a member of the Technical Committee on Multimedia Signal Processing of the IEEE Signal Processing Society and the Technical Committee on Visual Signal Processing and Communications of the IEEE Circuits and Systems Society. She has served on the organizing/technical committees of several international conferences and workshops, and as guest editor for several special issues related to image and

video coding. Her current research interests include image and video compression for unreliable networks, multimedia signal content analysis, motion estimation and object-oriented video coding, and inverse scattering for medical imaging.



Touradj Ebrahimi received his MSc and PhD degrees, both in electrical engineering, from the Swiss Federal Institute of Technology, Lausanne, Switzerland, in 1989 and 1992, respectively. From

1989 to 1992 he was a research assistant at the Signal Processing Laboratory of the Swiss Federal Institute of Technology (EPFL). During the summer of 1990 he was a visiting researcher at the Signal and Image Processing Institute of the University of Southern California, Los Angeles, California. In 1993 he was a research engineer at the Corporate Research Laboratories of Sony Corporation in Tokyo, where he conducted research on advanced video compression techniques for storage applications. In 1994 he served as a research consultant at AT&T Bell Laboratories working on very low bitrate video coding. He is currently at the Signal Processing Laboratory of EPFL, where he is involved with various aspects of digital video and multimedia applications and in charge of the digital TV group. In 1989 he was the recipient of the IEEE and Swiss National ASE award. His research interests are in multidimensional signal processing, image processing, information theory, and coding. He is the author or the co-author of more than 70 research publications and six patents. He is a member of IEEE and EURASIP.



Majid Rabbani received his PhD degree in electrical engineering from the University of Wisconsin in Madison in 1983 and joined the Eastman Kodak Research Laboratories in

Rochester, New York, that year. Currently, he is Senior Research Associate and heads the two technology areas of digital video processing and image compression within the Imaging Science Division of Kodak. He has actively represented Kodak at the JPEG International Committee since its inception over a decade ago. He is also involved in many educational activities, such as teach-

ing imaging short courses both at Kodak and outside, and an adjunct appointment at the Rochester Institute of Technology. He received the 1998 Kodak C.E.K. Meese Award (Kodak's highest research honor) and was co-recipient of two Engineering Emmy awards in 1990 and 1996. In 1992 he directed a team at Los Alamos National Labs to conduct the image analysis and enhancement of the videotape of the beating of Rodney King and subsequently testified as the prosecution expert witness in Federal Court in 1993. His current research interests span the various aspects of digital imaging where he has published over 50 technical articles and holds 13 patents. He is a Fellow of IEEE and SPIE, and is actively involved in organizing technical conferences and panels for both societies. He co-authored the book *Digital Image Compression Techniques* published by SPIE Press in 1991 and was editor of the SPIE Milestone Series

on *Image Coding and Compression*, published in 1992.

Ya-Qin Zhang received his BS and MS degrees in electrical engineering from the University of Science and Technology of China (USTC) in 1983 and 1985, respectively, and his PhD degree in electrical engineering from George Washington University, Washington D.C., in 1989. He is currently Director of the Multimedia Technology Laboratory at Sarnoff Corporation in Princeton, New Jersey (formerly David Sarnoff Research Center, and RCA Laboratories). His laboratory is a world leader in MPEG2/DTV, MPEG4/VLBR, and multimedia information technologies. He was with GTE Laboratories, Inc., in Waltham, Massachusetts and Contel Technology Center in Virginia from 1989 to 1994. He has authored and co-authored over 120 refereed papers and 20 US patents either pending or granted in image/video compression

and communications, multimedia, wireless networking, satellite communications, and medical imaging. He was an adjunct faculty member at George Washington University and Tufts University in 1990 and 1994. Dr. Zhang has been Editor-In-Chief for the *IEEE Transactions on Circuits and Systems for Video Technology* since July 1997 and was a co-Guest Editor for the Special Issue on Advances in Image and Video Compression of the *Proceedings of IEEE* in February 1995. He serves on the editorial boards of seven other professional journals and over 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 was honored as a Fellow of IEEE for his "contribution to and leadership in development of digital video compression and communications technologies, standards, and products."