GUEST EDITORIAL

Special Issues on Multimedia Communication Services

Emerging and established international standards such as MPEG-4, MPEG-7, H.263, and JPEG-2000 have spurred the development of multimedia services. The Internet, the world wide web, cellular phones, and affordable PCs/laptops have further contributed to this development. Multimedia (audio, text, graphics, images, video, animation, music) play a key role in web-based online services such as teleshopping, e-commerce, banking, stock market transactions, travel, video mail, sports, medicine, history, digital libraries, multimedia kiosks, and distance education/training. Indexing, browsing, and access/retrieval of multimedia in an efficient/effective way is no longer a novelty. Some of these aspects are being addressed in MPEG-7, "Multimedia Content Description Interface" whose primary objective is to facilitate the description, identification, and access of audiovisual data. As web-based online services proliferate, the privacy, security, and authenticity of multimedia transactions are essential. The need for high-quality, reliable, and inexpensive interactive multimedia services is clear.

After a thorough review/revision process, the following 16 papers have been accepted for the special issues. These papers have been grouped into two parts. The first seven papers constitute Part 1 and were published in CSSP issue no. 2, 2001. Part 2 contains the remaining nine papers and are included in this issue.

"Long Transition Analysis for Digital Video Sequences," by Wei Jyh Heng and King N. Ngan, introduces an automatic process that determines transition types and extracts information from them, for use in object extraction. Such a process consists of four stages: shot boundary refinement, shot type determination, frame reconstruction for soft transitions, and shot classification for hard transitions. The long transition analysis bridges the gap between shot boundary detection and object tracking, and smoothes the process of automatic video indexing for video databases.

"Segmentation of Moving Objects in Image Sequence: A Review," by Dengsheng Zhang and Guojun Lu, provides a review of this important and challenging area of spatial segmentation of moving objects. Common approaches including temporal segmentation, spatial segmentation, and the combination of temporal-spatial segmentation are described. As an example, a complete segmentation scheme, which is an informative part of MPEG- 4, is summarized.

"Edge-Preserving Disparity Estimation and Disparity-Compensated Inter-mediate View Reconstruction for Stereo Images," by Sung-Sik Kim, Jung-Young Son, Young Huh, Chulhee Lee, and Kwanghoon Sohn, proposes a constrained disparity estimation method that uses a directional regularization technique to efficiently preserve edges for stereo image coding. The proposed method smoothes disparity vectors in smooth regions and preserves edges in object boundaries well, without creating an oversmoothing problem. The experimental results show that the proposed disparity estimation method gives close matches between a left image and a right image and improves coding efficiency.

"Data Hiding in Ordered Dithered Halftone Images," by Ming Sun Fu and Oscar C. Au, proposes a novel method called data hiding ordered dithering (DHOD) to hide a relatively large amount of invisible watermarking data in ordered dithered halftone images while retaining good visual quality. Simulation results suggest that DHOD can indeed hide a large amount of data and still maintain good visual quality.

"Fast and Efficient Motion Estimation Using Diamond Zonal-Based Algo-rithms," by Alexis M. Tourapis, Oscar C. Au, Ming L. Liou, and Guobin Shen, proposes a novel algorithm called advanced diamond zonal search (ADZS), which was submitted to and well received by the Moving Pictures Experts Group (MPEG) standards committee for inclusion as an encoder optimization tool. ADZS was criticized by MPEG for using fixed thresholds, which may not be suitable for all video sequences. To address this issue, a threshold-adaptive version, called threshold-adaptive ADZS (TAADZS), is proposed. Simulation results verify the superior performance of ADZS and TAADZS over other fast algorithms and the robustness of TAADZS over ADZS.

"Performance of the Color Set Partitioning in Hierarchical Tree Scheme (CSPIHT) in Video Coding," by Ashraf A. Kassim and Lee Wei Siong, describes the implementation of the recently introduced color set partioning in hierarchical tree (CSPIHT)-based scheme for video coding. The intra- and interframe coding performance of a CSPIHT-based video coder (CVC) is compared against that of the H.263 at bit rates lower than 64 kbit/s. The CVC performs comparably or better than the H.263 at lower bit rates, whereas the H.263 performs better than the CVC at higher bit rates.

"A Logical Framework for Visual Information Modeling and Management," by Youngchoon Park, Pan Koo Kim, Forouzan Golshani, and Sethuraman Pan-chanathan, presents a unified semantic visual data- modeling framework. An ex-tended conceptual graph is proposed as an annotation mechanism of a user's understanding of video objects, activities, and events. The proposed visual data model has six different abstraction layers. A higher level is more abstracted and more semantically summarized. A polygon-based bounding volume is used in video object approximation in space and time. A bounding volume in motion

trajectory representation is used, rather than motion vectors. This model may be used as a referencing framework for various visual information management systems' developments.

"Performance Comparison of MPEG-4 and H.263+ for Streaming Video Applications," by Krit Panusopone and Ajay Luthra, compares the video coding performance of both MPEG-4 and H.263+ standards for delivering streaming video over the Internet. It also highlights the appropriate combinations of the tools (MPEG-4) and the options (H.263+) that provide good performance for streaming video applications.

"Audio-Visual Integration in Multimedia Communications Based on MPEG-4 Facial Animation," by Z. S. Bojkovic and D. A. Milovanovic, reviews coding methods for bit rate reduction of facial animation parameters, which make possible the transmission of multiple talking heads over band-limited channels. Further, relationships between natural/synthetic audio/video coding from the point of view of integration of face animation with natural video are emphasized. Within MPEG-4, a binary format for scene (BIFS) description framework offers a parametric methodology for scene structure representation and efficient coding for transmission or storage. The MPEG-4 profiling strategy in facial animation, which guarantees that the standard can provide adequate solutions for applications in multimedia communications, is addressed.

"Online Traffic Smoothing for Delivery of VBR Media Streams," by Ray-I Chang, Meng-Chang Chen, Jan-Ming Ho, and Ming-Tat Ko, proposes a new window-based method for online traffic rooting for delivery of VBR media streams. It introduces two new ideas, the dynamic window-sliding size and the aggressive workahead, for delivery of online VBR media streams. The aggressive and dynamic window-sliding (ADWS) method can automatically decide the suitable window-sliding sizes for different windows. Thus, the allocated peak bandwidth can be further reduced. By examining various media streams, ADWS is shown to be effective and efficient.

"Online Rate Control for Video Streams," by Sassan Pejhan, Tihao Chiang, and Ya-Qin Zhang, describes a mechanism for varying the frame rate of re-encoded video clips online. The mechanism relies on two different encoders. An offline encoder creates a high-quality bit stream encoded at 30 fps, as well as separate files containing motion vectors for the same clip at lower frame rates. An online encoder decodes the bit stream (if necessary) and re-encodes it at lower frame rates in real time using the precomputed, stored motion information.

"Synthesis of Resources Sharing," by Y. Q. Zhang, C. S. Choy, and C. F. Chan, presents an algorithm of DSP processor design with high throughput and low cost by data pipelining. The hardware resources, which are composed of function units (FUs), register units (RUs), bus units (BUs), and memory units (MUs) in an executing model, are described. Under the constraints in the library, the DSP

data was read in as a control data flow graph (CDFG), the resources selection, mapping, and sharing were conducted based on this algorithm.

"Universal Multimedia Access from Wired and Wireless Systems," by A. Perkis, Y. Abdeljaoued, C. Christopoulos, T. Ebrahimi, and J. Chicharo, discusses issues with regard to enabling terminals of limited communications, processing, storage, and display capabilities to access rich multimedia contents anytime and anywhere. The universal multimedia access (UMA) concept described by the authors may provide a framework that will impact on the future development of personal computing and communication systems and devices.

"Modeling and Prediction of Hybrid Coded VBR Video Sources in Fuzzy Logic Perspectives," by B. Qiu, proposes a novel fuzzy logic prediction method that is suited to fast computation for online operation and has better prediction performance in terms of the error mean and the standard deviation than the autoregressive prediction. It can be used in the design of connection admission control, usage parameter control, and congestion control algorithms for multimedia communication networks.

"On the Importance of Error Resilience in Visual Communications over Noise Channels," by A. Perkis, addresses an important issue in digital image transmission. It provides a detailed review of the field and may serve well as an introduction for readers who are new to the field. The paper investigates three error resilience schemes, which include substitution of the quantization and symbol encoding by a fixed-length coding scheme, substitution by a mixed fixed-length coding and variable-length coding, and substitution of the variable-length coding by a reversible variable-length coding.

"A New Efficient Expression Generation and Automatic Cloning Method for Multimedia Actors," by S. Karunaratne and H. Yan, presents a very comprehensive survey of the state of the art in the field of face animation. The paper describes a new method for facial expression generation on cloned synthetic head models, which is shown to be effective compared with a number of existing techniques. It has applications in multimedia communications, educational agents, and synthetic actors in movies and games.

Professor Hsueh-Ming Hang
National Chiao Tung University
Center for Telecommunications Research GmbH
1001 Ta-Hsueh Road
Hsin-chu, 300-50 Taiwan
Fax: 886-35-723283
E-mail: hmhang@cc.nctu.edu.tw

Dr. Kamisetty R. Rao
Professor of Electrical Engineering
416 Yates St., Rm 518/Box 19016
University of Texas at Arlington
Arlington, TX 76019-0016
USA
Tel: +817-272-3478
Fax:+817-272-2253
E-mail: krrao@exchange.uta.edu

Dr. Thomas Sikora
Image Processing Department
Heinrich-Hertz Institute (HHI) Berlin
Einsteinufer 37, 10587 Berlin
Germany
Tel: +49-30-31002-622,
Fax: +49-30-392-7200
E-mail: sikora@hhi.de

Dr. Hong Ren Wu
Associate Professor
School of Computer Science &
Software Engineering
Monash University, Wellington Road
Clayton 3168, Australia
Tel: 61-3-9905-3255
Fax: 61-3-9905-5146
E-mail: hrw@dgs.monash.edu.au