Introduction to the Special Issue on the H.264/AVC Video Coding Standard

OVER THE LAST one and a half decades, digital video and audio compression technologies have fundamentally changed the way we create, communicate, and consume audio-visual information. They have not only transformed existing applications and services like the distribution of entertainment video to the home but also spawned brand new industries and services like video-conferencing, direct-to-home satellite distribution, digital video recording, video-on-demand services, high-definition TV, video on mobile devices, streaming video, etc.

Since the early 1990s, when the technology was in its infancy, international video coding standards such as H.261, MPEG-1, MPEG-2/H.262, H.263, and MPEG-4 (Part 2) have been powerful engines behind the commercial success of digital video compression. They have played pivotal roles in establishing the technology via providing interoperability among products developed by different manufacturers and at the same time allowing flexibility for ingenuity in optimizing and molding the technology to fit a given application and making the cost-performance tradeoffs well suited to particular requirements. They have provided much needed assurance to the content creators that their content will run everywhere and that they do not have to create and manage multiple copies of the same content to match the products of different manufacturers. They have permitted the economy of scale to allow steep reduction in cost for masses to be able to afford the technology. They have nurtured open interactions among experts from different companies to spurt exponential growth in innovation to keep pace with the implementation technology and the needs of the applications.

The subject of this Special Issue is the newest entry in the line of international video coding standards, referred to formally as ITU-T Recommendation H.264 and ISO/IEC MPEG-4 (Part 10) Advanced Video Coding (or referred to in short as H.264/AVC). H.264/AVC is the powerful and state-of-the-art video compression standard recently developed by the ITU-T/ISO/IEC Joint Video Team (JVT) consisting of experts from ITU-T's Video Coding Experts Group (VCEG) and ISO/IEC's Moving Picture Experts Group (MPEG). As has been the case with past video coding standards, its design represents a delicate balance between coding gain, implementation complexity, and costs based on state of VLSI (ASICs and Microprocessors) design technology. It has brought the coding algorithm technology in synch with the VLSI technology that is available today or will be available in near future. Some algorithmic features that could not be considered in the past due to implementation complexity could now be accepted. New techniques developed since the time of the prior designs could be incorporated. Some features that provided some increase in coding gain but were too expensive to implement were not included. In the process, the H.264/AVC design emerged with an improvement in coding efficiency typically by a factor of two over MPEG-2—the most widely used video coding standard today—while keeping the cost within the acceptable range. Similar quality gains were shown relative to other prior designs for a very broad range of applications. Above all, this project has proven that the predictions of the demise of video compression algorithm innovations were little premature and that they are continuing to flourish vigorously.

This Special Issue opens with an overview paper about the H.264/AVC standard by Wiegand *et al.* As the title suggests, it provides a concise description of the most important tools used in the standard.

We then dive deeper into the algorithm. We start with two papers on inter-picture prediction and processing. Wedi and Musmann provide good insight into the workings of fractional-sample motion-compensated prediction in their paper "Motion and Aliasing Compensated Prediction for Hybrid Video Coding". Flierl and Girod describe how and why B-pictures work and some ways the B picture concept has been extended in H.264/AVC in their paper "Generalized B Pictures and the Draft H.264/AVC Video Compression Standard".

The next two papers focus on the transforms used and considered in the standardization process. The first of these, "Low-Complexity Transform and Quantization in H.264/AVC" by Malvar *et al.*, describes the integer DCT-like transform and quantization scaling methods adopted in the standard. This transform was designed with particular attention paid to the fact that real-number-based transform specification causes mismatches between encoders and decoders, and can be expensive to implement in products that do not have high end microprocessor or ASIC. The second paper, "Variable Block Size Transforms for H.264/AVC" by Wien, describes the effort spent by JVT on adapting the transform size to the motion estimation block size. This technology was considered to be very interesting but not fully mature or sufficiently clearly beneficial to coding efficiency, and a bit too difficult to implement for inclusion in the standard at this stage.

Deblocking filtering has been one of the centers of great discussions among the experts and has divided the video compression community for a long time. In H.264/AVC, years of past experience in this field have culminated into the development of a new and powerful design. In their paper "Adaptive Deblocking Filter" List *et al.* explain how the filter adapts itself based on the quantization parameter, different compression modes, and motion in a scene to be able to differentiate compression artifacts from the content of the scene. The result has been a filter design that has reasonable implementation complexity and provides a substantial benefit to both subjective and objective video quality.

Digital Object Identifier 10.1109/TCSVT.2003.815169

Entropy coding of H.264/AVC syntax elements using "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard" is described by Marpe *et al.* Their paper provides information about the design for context-based adaptive binary arithmetic coding (CABAC) and provides background about various decisions made during the development of CABAC.

Karczewicz and Kurçeren explain the technology behind two new types of slices in H.264/AVC, which are SP and SI. Their paper "The SP and SI Frames Design for H.264/AVC" provides the concepts behind this new technology and also discusses how it can be used in various applications like streaming video, etc.

Two of the many key applications that the stanadardization committee had in mind are the Internet and wireless-based distribution of the video. Both of these media provide a challenge due to, among other things, highly error-prone transmission channels. The paper by Wenger entitled "H.264/AVC over IP" provides in-depth understanding of the tools included in H.264/AVC targeted specifically to address that challenge. The paper "H.264/AVC in Wireless Environments" by Stockhammer *et al.*discusses normative and nonnormative error-resilience extensions at the encoder and decoder when using H.264/AVC in error-prone wireless environments.

Interoperability is one of the key benefits provided by standards. To achieve interoperability, especially for real-time decoding, it is not sufficient to only specify an algorithm. It is important to also provide a description for how the bits arrive and what buffer models and memory resources an encoder assumes for the decoder. Ribas-Corbera *et al.* present the novel buffer model concept used in the standard in their paper "A Generalized Hypothetical Reference Decoder for H.264/AVC".

H.264/AVC standardizes decoder function, leaving room for encoder optimization. In the paper "Rate-Constrained Coder Control and Comparison of Video Coding Standards," Wiegand *et al.* present some of the ways to optimize an encoder in terms of coding efficiency. They apply the same technique to H.264/AVC, as well as to various preceding standards, and in this way provide a comparison between them on that basis.

The improvement in coding efficiency by H.264/AVC comes with a certain degree of complexity. This has been analyzed in two contributions—one by Horowitz *et al.*entitled "H.264/AVC Baseline Decoder Complexity Analysis" and one by Lappalainen *et al.* in their paper "Complexity of Optimized H.264/AVC Video Decoder Implementation".

We hope that this Special Issue provides a useful perspective on the core coding tools used in H.264/AVC. We would like to thank the authors for their excellent contributions. The standard moved almost at light speed and the authors did a fabulous job of keeping pace with it and updated their manuscripts more than once. We believe the end result is a good balance between providing enlightening information to the community in a timely fashion and providing a complete portrait of the concepts and content of the final standard. We are also deeply indebted to the reviewers for their quick and very constructive responses.

Moreover, we are grateful to all those companies and universities who participated in the development of this standard. They sent some of their brightest minds around the world to produce this standard. Our hats off to all the experts who were involved in giving birth to this standard for tirelessly attending the meetings—some of them lasted nine continuous days with some of those days closing between 2 to 6 A.M. to start again at 8 or 9 A.M., and for spending countless hours going through hundreds of pages and thousands of software lines to make sure that the text and code are accurate and embody the best possible design within the limits of our collective abilities.

Finally, we would also like to thank W. Li, former Editor-in-Chief, for suggesting this Special Issue, and T. Sikora, current Editor-in-Chief, for his patience, encouragement, and support for pursuing it.

AJAY LUTHRA
Broadband Communications Sector
Motorola, Inc.
San Diego, CA 92121 USA

GARY J. SULLIVAN
eHome A/V Platforms Group
Microsoft Corporation
Redmond, WA 98052 USA

THOMAS WIEGAND
Fraunhofer–Institute for Telecommunications
Heinrich–Hertz Institue
10587 Berlin, Germany

Ajay K. Luthra (S'79–M'81–SM'89) received the B.E. (Hons.) degree from BITS, Pilani, India, in 1975, the M.Tech. degree in communications engineering from IIT Delhi, Delhi, India, in 1977, and the Ph.D. degree from Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, in 1981.

From 1981 to 1984, he was a Senior Engineer at Interspec Inc., Philadelphia, PA, where he was involved in applications of digital signal and image processing for bio-medical applications. From 1984 to 1995, he was with Tektronix, Beaverton, OR, where he was Manager of the Digital Signal and Picture Processing Group during 1985–1990 and then Director of the Communications/Video Systems Research Lab from 1990–1995 . He is currently a Senior Director in Advanced Technology Group at Motorola, Broadband Communications Sector (formerly General Instrument), San Diego, CA, where he is involved in advanced development work in the areas of digital video compression and processing, streaming video, interactive TV, cable head-end system design, and advanced set-top box architectures. He has been an active member of the MPEG committee for the last ten years where he has chaired several technical subgroups. He is an Associate Rapporteur/Co-Chair of the Joint Video Team consisting of ISO/MPEG and ITU-T/VCEG experts working on developing next generation of video coding standards.

Dr. Luthra was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (2000–2002) and also a Guest Editor for its Special Issue on Streaming Video (March 2001).

Gary J. Sullivan (S'83–M' 91–SM'01) received the B.S. and M.Eng. degrees in electrical engineering from the University of Louisville, Louisville, KY, in 1982 and 1983, respectively, and the Ph.D. and Eng. degrees in electrical engineering from the University of California, Los Angeles, in 1991.

He is the Chairman of the Joint Video Team (JVT) for the development of the next-generation H.264/MPEG4-AVC video coding standard, which was recently completed as a result of a joint project between the ITU-T video coding experts group (VCEG) and the ISO/IEC moving picture experts group (MPEG). He is also the Rapporteur of Advanced Video Coding in the ITU-T, where he has led VCEG (ITU-T Q.6/SG16) for about six years, and the ITU-T Video Liaison Representative to MPEG (ISO/IEC JTC1/SC29/WG11) and served as MPEG's Video Chairman during 2001–2002. He is currently a Program Manager of video standards and technologies in the eHome A/V Platforms Group of Microsoft Corporation, Redmond, WA, where he designed and remains lead engineer for the DirectX® Video Acceleration API/DDI feature of the Microsoft Windows® operating system platform. Prior to joining Microsoft in 1999, he was the Manager of Communications Core Research at PictureTel Corporation, the quondam world leader in videoconferencing communication. He was previously a Howard Hughes Fellow and Member of the Technical Staff in the Advanced Systems Division of Hughes Aircraft Corporation and was a terrain-following radar system software engineer for Texas Instruments. His research interests and areas of publication include image and video compression, rate-distortion optimization, motion representation, scalar and vector quantization, and error and packet loss resilient video coding.

Thomas Wiegand received the Dr.-Ing. degree from the University of Erlangen-Nuremberg, Germany, in 2000 and the Dipl.-Ing. degree in electrical engineering from the Technical University of Hamburg-Harburg, Germany, in 1995.

He is the Head of the Image Communication Group in the Image Processing Department, Fraunhofer–Institute for Telecommunications –Heinrich Hertz Institute (HHI), Berlin, Germany. During 1997 to 1998, he was a Visiting Researcher at Stanford University, Stanford, CA, and served as a Consultant to 8x8, Inc., Santa Clara, CA. From 1993 to 1994, he was a Visiting Researcher at Kobe University, Kobe, Japan. In 1995, he was a Visiting Scholar at the University of California at Santa Barbara, where he began his research on video compression and transmission. Since then, he has published several conference and journal papers on the subject and has contributed successfully to the ITU-T Video Coding Experts Group (ITU-T SG16 Q.6—VCEG)/ISO/IEC Moving Pictures Experts Group (ISO/IEC JTC1/SC29/WG11—MPEG)/Joint Video Team (JVT) standardization efforts and holds various international patents in this field. He has been appointed as the Associated Rapporteur of the ITU-T VCEG (October 2000), the Associated Rapporteur/Co-Chair of the JVT that has been created by ITU-T VCEG and ISO/IEC MPEG for finalization of the H.264/AVC video coding standard (December 2001), and the Editor of the H.264/AVC video coding standard (February 2002).