The H.264/MPEG-4 AVC video coding standard and its deployment status

Gary J. Sullivan

Microsoft Corporation, One Microsoft Way, Redmond, WA 98052, USA

ABSTRACT

The new video coding standard known as H.264/MPEG-4 Advanced Video Coding (AVC), now in its fourth version, has demonstrated significant achievements in terms of coding efficiency, robustness to a variety of network channels and conditions, and breadth of application. The recent fidelity range extensions have further improved compression quality and further broadened the range of applications, and the recent corrigenda have excised the inevitable errata of the initially-approved versions of the specification. Patent licensing programs have begun, the standard has been adopted into a variety of application specifications, and products suitable for widespread deployment have begun to appear. New work toward the near-term development of scalable video coding (SVC) extensions is also under way. This paper does not attempt to review the details of the H.264/MPEG-4 AVC technical design, as that subject has been covered already in a number of publications. Instead, it covers only the high-level design characteristics and focuses more on the recent developments in the standardization community and the deployment status of the specification.

Keywords: H.264, MPEG-4, Advanced Video Coding (AVC), video coding, standards, VCEG, MPEG, JVT, H.26L.

1. INTRODUCTION

The latest international standard for video coding, known as H.264/MPEG-4 Advanced Video Coding (AVC)[1], is the newest entry in the series of international video coding standards, and it has begun a strong transition from a specification development project to widespread product deployment.

It is currently the most powerful and state-of-the-art such standard, and was developed by a Joint Video Team (JVT) consisting of experts from two "parent bodies": the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG). The design of the new standard provides the most current balance between coding efficiency, error/loss robustness, and implementation complexity, – based on the current state of VLSI design technology (CPU's, DSP's, ASIC's, FPGA's, etc.). In the process, a standard was created that improved coding efficiency by a factor of at least about two (on average) over MPEG-2 video[2] – the most widely used video coding standard today – while keeping the cost within an acceptable range for near-term deployment. In July 2004, a new amendment was added to this standard, called the *fidelity range extensions* (FRExt, Amendment 1), which demonstrates even further coding efficiency, particularly for high-resolution video applications such as HDTV.

A number of technical overviews and tutorials on the standard have been published (e.g., references 3-11, including two books on the subject[9, 10]), as well as a large number of publications on subjects of more narrow focus, such as studies of the design and optimization of particular parts of the specification (e.g., reference 12 and various papers in reference 13) or studies of the coding efficiency (e.g., reference 14) or implementation complexity of the new standard (e.g., references 7, and 15-19), or its use in specific environments (e.g., references 20 and 21). Most published overviews of the standard include only the scope of its original version. Among the overviews and tutorials, some notable treatments include reference 4 for information about the recent enhancements of the standard (the fidelity range extensions), references 3, 8, and 9 for their tutorial approach to the subject, references 5, 6, 9, and 10 for providing a relatively large amount of technical detail, and reference 7 for providing more analysis of implementation complexity issues.

This paper does not attempt to review the details of the H.264/MPEG-4 AVC technical design, as that subject has been covered already in a number of publications. Instead, it covers only the high-level design characteristics and focuses more on the recent developments in the standardization community and the deployment status of the specification. Section 2 covers the standardization development work, including recent efforts on fidelity range extensions, scalable video coding, corrigendum activity, and the development of related specifications. Section 3 discusses the high-level

Visual Communications and Image Processing 2005, edited by Shipeng Li,
Fernando Pereira, Heung-Yeung Shum, Andrew G. Tescher, Proc. of SPIE Vol. 5960
(SPIE, Bellingham, WA, 2005) · 0277-786X/05/$15 · doi: 10.1117/12.631573

technical characteristics of the design. Section 4 covers patent licensing considerations, and Section 5 discusses deployments and adoptions of the new standard. Section 6 offers conclusions on the current state of the standard.

2. STANDARDIZATION DEVELOPMENT WORK

2.1 Initial specification development background

The initial version of H.264/AVC was developed over a period of about four years. The roots of this standard lie in the ITU-T's H.26L project initiated by the Video Coding Experts Group (VCEG), which issued a Call for Proposals (CfP) in early 1998 and created a first draft design for its new standard in August of 1999. In 2001, when ISO/IEC's Moving Pictures Experts Group (MPEG) had finished development of its then-most recent video coding standard, known as MPEG-4 Part 2 video[23], it issued a similar CfP to invite new contributions to further improve the coding efficiency beyond what was achieved on that project. VCEG chose to provide its draft design in response to MPEG's CfP and proposed joining forces to complete the work. Several other proposals were also submitted and were tested by MPEG as well. The ITU-T's H.26L was a top-performing proposal, and most others that showed good performance in MPEG had also been based on H.26L (as it had become well-known as an advance in technology by that time). As a result of those tests, MPEG affirmed the design choices made by VCEG for H.26L, and the two groups joined forces to complete the work.

Therefore, to allow speedy progress, ITU-T and ISO/IEC agreed to join forces together to jointly develop the next generation of video coding standard and use H.26L as the starting point. A Joint Video Team (JVT), consisting of experts from VCEG and MPEG, was formed in December 2001 and it completed the technical development of the standard in May of 2003. As further encouragement for open and rapid progress, nearly all documents involved in the work of the JVT were made publicly available throughout the development of the standard[24] (similarly, the documents of VCEG, where the H.26L project previously was found, have also been available[25]). The ITU-T approved the standard under its naming structure as ITU-T Rec. H.264, and ISO/IEC approved it as ISO/IEC 14496-10 Advanced Video Coding (AVC), in the "MPEG-4" suite of standards. As an unwanted byproduct, this standard sometimes gets referred to by at least six different names – H.264, H.26L, ISO/IEC 14496-10, MPEG-4 AVC, MPEG-4 Part 10, and JVT. Using mixtures of these names has become typical for clarity and balance, so in this paper we refer to it as H.264/MPEG-4 AVC to strike a balance between names used by the two partnered organizations.

With the wide breadth of applications considered by the two organizations, the application focus for the work was correspondingly broad – from videoconferencing to entertainment (including broadcasting over cable, satellite, terrestrial, cable modem, DSL, etc.; storage on DVDs and hard disks; video on demand etc.) to streaming video, surveillance and military applications, and digital cinema. Three basic feature sets called *profiles* were established to address these application domains:

– **Baseline profile (BP)**: designed to minimize complexity and provide high robustness and flexibility for use over a broad range of network environments and conditions;
– **Main profile (MP)**: designed with an emphasis on compression *coding efficiency* capability; and
– **Extended profile (XP)**: designed to combine the robustness of the Baseline profile with a higher degree of coding efficiency and greater network robustness, and to add enhanced modes useful for special "trick uses" for such applications as flexible video streaming.

Among these, implementation effort has thus far focused primarily on the Baseline and Main profiles. At the moment, the Baseline profile appears to provide a good solution for its target application area, but the primary interest that initially focused on the Main profile has now shifted to the High profile of FRExt.

The standard also defines a large number (currently sixteen) of distinct "levels" of decoder capability, ranging from the most basic "level 1" having only the capability to decode video with a maximum luma resolution of 176x144 at 15 frames per second with a compressed bit rate of approximately 64 kbps, on up to "level 4" and "level 4.1" which have HDTV capability and three more levels beyond that (up to "level 5.1" for 4096x2048 video at 30 frames per second or 4096x2304 video at 26.7 frames per second with compressed bit rates in the hundreds of Mbps).

2.2 Fidelity Range Extensions (FRExt)

While having a broad range of applications, the initial H.264/MPEG-4 AVC standard (as it was completed in May of 2003), was primarily focused on "consumer-quality" video, based on SDTV or lower video resolution, 8-bits/sample, and 4:2:0 chroma sampling. Due to the committee's schedule goals, it did not include support for use in the most demanding professional environments, and the design had not been focused on the highest video resolutions. To address the needs of these most-demanding applications, a continuation of the joint project was launched to add new extensions to the capabilities of the original standard. This effort took about one year to complete – starting with a first draft in May of 2003, the final specification was completed in July of 2004 and the editing period was completed in August of 2004. These extensions, originally known as the "professional" extensions, were eventually named the "fidelity range extensions" (FRExt) to better indicate the spirit of the effort. One overview paper[4] is notable in regard to this topic, as it includes a description of the FRExt enhancement features.

In the process of designing the FRExt amendment, the JVT was also able to go back and re-examine some prior technical proposals that had not been included in the initial standard due to scheduling constraints, uncertainty about benefits, or the original scope of intended applications. With the additional time afforded by the extension project, it was possible to include some of those features in the new extensions. The FRExt project produced a suite of four new profiles collectively called the *High* or FRExt profiles:

- **High profile (HP)**: supporting 8-bit video with 4:2:0 sampling, addressing high-end consumer use and other applications using high-resolution video without a need for extended chroma formats or extended sample accuracy,
- **High 10 profile (Hi10P)**: supporting 4:2:0 video with up to 10 bits of representation accuracy per sample,
- **High 4:2:2 profile (H422P)**: supporting up to 4:2:2 chroma sampling and up to 10 bits per sample, and
- **High 4:4:4 profile (H444P)**: supporting up to 4:4:4 chroma sampling, up to 12 bits per sample, and additionally supporting efficient lossless region coding and an integer residual color transform for coding RGB video while avoiding color-space transformation error.

All of these profiles support all features of the prior Main profile, and additionally support an adaptive transform block-size and perceptually-optimized encoder-selected quantization scaling matrices.

Initial industry feedback has been dramatic in its rapid embrace of FRExt. The High profile, in particular, appears certain to be incorporated into several important near-term application specifications and deployments.

2.3 Corrigendum work and versions of the standard

Despite strong effort by its developers, nearly every complex detailed technical specification will initially contain errors and need follow-up work to correct them. After developing the first version of the standard, the JVT found and fixed a substantial number of such problems in a second version of the standard produced in March of 2004. Recently, after completing the FRExt amendment (resulting in what the JVT calls version 3), another round of corrections has been completed. The current version of the standard is referred to by the JVT as version 4. By the time of this conference, a version 4 text (on which the drafting work was approximately completed near the end of 2004) should be completed for both the ITU-T and ISO/IEC.

2.4 Scalable video coding (SVC) extensions

The JVT is presently working on the development of a new set of extensions of the capability of H.264/MPEG-4 AVC. The new work addresses a desire to incorporate a *scalable video coding* (SVC) amendment into the design in a maximally-compatible way. Here the sense of the term *scalability* is to refer to a functionality that allows the removal of parts of the bitstream while achieving a rate-distortion performance with the remaining data (at any supported spatial, temporal, or SNR resolution) that is comparable to "single-layer" H.264/MPEG-4 AVC coding (at that particular resolution). The intent is to finish the drafting work on the new extensions by July of 2006. Here, "comparable" is taken to refer to achieving coding efficiency within about 10% excess bit rate for the same decoded video fidelity.

The scalability extensions will build on H.264/MPEG-4 AVC in a backwards-compatible way. In terms of coding structure, a scalable bitstream will be composed of a *base layer* bitstream and one or several *enhancement layer* bitstreams. The base layer will be conforming to one of the profiles of the prior H.264/MPEG-4 AVC design. A single-layer H.264/MPEG-4 AVC decoder (as currently specified) will be capable of decoding the base layer by ignoring the parts of the bitstream that correspond to the enhancement layer(s) according to data insertion methods already specified in the standard. For the enhancement layer specification, the design will also be done in a maximally-consistent fashion – any change to the current H.264/MPEG-4 AVC syntax and decoding process must be well-justified by sufficient relative rate-distortion improvements.

Some forms of scalability that will be supported include spatial scalability (e.g., a base layer bitstream that represents QCIF video and an enhancement layer bitstream that represents CIF video), quality ("SNR") scalability, and temporal scalability (something already supported in to a large extent by H.264/MPEG-4 AVC as it exists; however, there may be a need to enhance the quality of the video achieved with temporal scalability functionality in order to achieve the best practical quality). Combined scalability structures will also be included, and roughly similar characteristics to the current design should be achieved in terms of achievable complexity, end-to-end delay characteristics, and robustness to transmission errors and congestion.

2.5 Conformance and reference software standardization

Two additional standards were developed to directly support the video coding specification itself. As with the video coding standard, these were developed and standardized jointly in the ITU-T and ISO/IEC.

The first of these is the development of tests for conformance to the video coding standard. A set of test bitstreams and procedures for conformance testing were jointly developed by the JVT and standardized by the parent bodies as ITU-T Rec. H.264.1 and as amendments to ISO/IEC 14496-4 (MPEG-4 Conformance).

The other is the development of reference software. The reference software development process has been managed jointly through a publicly-accessible web site[26], and the software is standardized by the parent bodies as ITU-T Rec. H.264.2 and amendments to ISO/IEC 14496-5 (MPEG-4 Reference Software).

3. TECHNICAL DESIGN CHARACTERISTICS

3.1 Coding efficiency

A major goal (probably *the primary* major goal) of the development of the new standard was to enable a major improvement in *coding efficiency*, a term which refers to the number of compressed video bits necessary to obtain a particular decoded video fidelity.

The coding efficiency associated with using a video coding specification is a very difficult thing to measure, since many considerations can interfere with attempts to measure it, including factors such as the following:

– The lack of objective performance measures that match the human visual system perception of video distortion.
– The time, energy, and expense of tests involving human perceptual measurements.
– The imperfections that creep into tests involving human subjects.
– The statistical random error involved in test measurements of subjective quality.
– The difficulty of determining the best way to use a standardized syntax to maximize quality.
– The difficulties of determining the reference design(s) against which to compare.
– The difficulty of finding test video clips that represent the corpus of content that is relevant to a target application.
– The wide breadth of applications intended to be covered by a standardized specification.
– The ability of decoders to use (proprietary or standardized) post-processing, scaling, and display customization.
– The ability of encoders to perform (proprietary or standardized) pre-processing before encoding the video content.
– The variety of profiles and levels specified in the standard.

Indeed, MPEG-2 usage experience over the last ten years or so has shown that the quality that is achievable in products using a particular standardized syntax can continue to improve over significant periods of time (and in a difficult-to-predict fashion) as industry learns how to make better and better products.

It should be noted that the scope of a video coding standard is intentionally limited – to allow maximal encoder and decoder design freedom within the bounds of what is necessary to achieve full *interoperability*. This scope restriction ensures that all decoders that implement a profile and level can decode all video that is intended for that configuration, but provides no guarantees of the quality that will be produced by encoders. Indeed, even crude encoding techniques are allowed to be used with claims of full conformance to the standard.

Despite the difficulty, the development and deployment of video coding standards requires best-effort expert judgments to be made about what coding efficiency is achievable in practice and how that efficiency will evolve over time. During the development of the standard, controlled tests using reference software and based on combinations of objective performance measures (such as peak signal-to-noise-ratio PSNR) and formal or informal subjective quality assessments were performed to design the syntax and decoding process for the standard.

Upon completion of the first version of the standard, a substantial *verification testing* effort was conducted by the MPEG parent body to assess its coding efficiency capability for a variety of application environments. The tests were largely based on the well-established ITU-R BT.500-11 formal subjective testing methodology[27]. The results of that study were made public[28], and additional publications relating to that study have also appeared[29, 30]. These tests evaluated the performance of H.264/MPEG-4 AVC as compared to the MPEG-4 Part 2 Visual[23] (ISO/IEC 14496-2) and MPEG-2 Video[2] (ISO/IEC 13818-2 | ITU-T Rec. H.262) standards. The overall results showed that H.264/MPEG-4 AVC achieved a coding efficiency improvement of 1.5 times or greater in 78% (66 out of 85) of the statistically conclusive cases, out of which 77% (51 out of 66) showed improvements of 2 times or greater, relative to its predecessors. It was concluded that H.264/MPEG-4 AVC could deliver acceptable or even good quality at bit rates as low as 1.5 Mbps and 6 Mbps for SD and HD sequences, respectively – these being bit rates at which MPEG-2 clearly could not deliver acceptable video quality. It was further remarked that there was much space for further optimization and that future encoders probably will deliver broadcast quality at bit rates even lower than what was selected for the test. This test included only the original version of the standard.

Other subjective studies have been done for use in particular application environments. Some results of one such study for the new FRExt High profile were made available[31], which included the results shown in Fig. 1 below. This test exhibited a nominal 3:1 coding efficiency gain for H.264/MPEG-4 AVC relative to MPEG-2 (with 8 Mbps use of H.264/MPEG-4 AVC High profile scoring slightly higher on average than MPEG-2 at three times the bit rate) and achieving approximate transparency at 16 Mbps for 1920x1080 progressive scan film material at 24 frames per second.

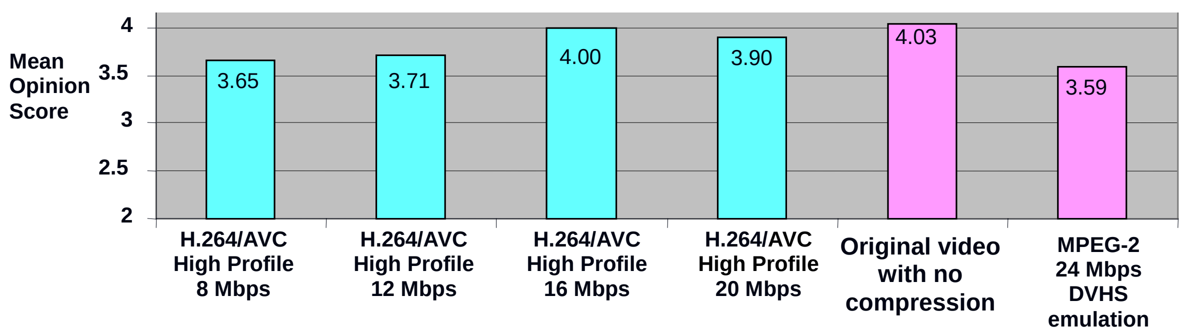

Figure 1. Subjective test results for H.264/MPEG-4 AVC High Profile and MPEG-2 Main Profile (from reference 31)

Using objective testing methodologies (esp., PSNR), a number of studies have been performed – and with roughly similar results. Descriptions of some such studies are found in some of the overview papers. One notable such study used well-documented consistently-designed optimization methods to compare the quality for various standards[14]. Table 1, from that study, shows the estimated bit rate savings of H.264/MPEG-4 AVC relative to key profiles of three prior standards, for a set of popular lower-resolution *quarter common intermediate format* (QCIF) (10 Hz and 15 Hz) and somewhat-higher-resolution *common intermediate format* (CIF) (15 Hz and 30 Hz) video sequences with different amounts of motion and different spatial detail characteristics.

Table 1. Measured bit rate savings of H.264/MPEG-4 AVC relative to key profiles of prior standards (from reference 14)

Coder	MPEG-4 ASP	H.263 HLP	MPEG-2 MP
H.264/MPEG-4 AVC	39%	49%	65%

Most experts seem to agree that the new standard can achieve roughly double the coding efficiency (i.e., a 50% reduction in the number of bits needed for a particular level of fidelity) of prior standards (particularly relative to MPEG-2 video), or perhaps even more, particularly when the FRExt High profile is included in the analysis.

3.2 Network robustness

Another key goal in the development of the standard was to achieve a high degree of flexibility for application to uses in many networks and storage-based environments. This is reflected in a number of aspects of the design, perhaps most notably in the structuring of the standard in terms of a *video coding layer* (VCL) responsible for efficient representation of video content and a *network abstraction layer* (NAL) responsible for structuring the coded video data into a form suitable for use in a wide variety of environments. In addition to this basic structuring, a number of specific technical features were put into the design to achieve robustness to difficult network environments, including losses of data due to congestion and network errors. Some published studies[20, 21] have focused on the network robustness behavior.

3.3 Implementation complexity

Often, when designing algorithms, there is a tradeoff between performance and complexity. The goal during the development of the new standard was to maximize quality without making the implementation of the design impractical (particularly for decoders, as encoder technology can evolve over time but the decoding process must be defined precisely and must be sufficiently practical for the initial deployments). Moore's law has helped, to the extent that some algorithmic techniques that would previously have been considered impractical are now no longer especially difficult (although video tends to always be somewhat of a challenge). Here we use the term "complexity" to refer to the amount of computational resources needed (in terms of memory capacity for instructions and data, bus bandwidth, instructions per second, cache-access issues, pipelined and SIMD processing considerations, etc.) needed for an implementation.

Architecture dependencies make measurements of design complexity difficult. Additionally, a significant amount of freedom is allowed within the bounds of the standard. Decoders are allowed to compute the decoded result in any way they choose as long as they compute the correct values for the decoded samples of every picture (and do so at the necessary speed). Encoders are allowed to be designed in any way that does not result in bitstreams that violate the syntax constraints specified in the standard. And the attempt to cover a very wide breadth of applications for the new standard makes it difficult to ensure that the complexity bounds of each application environment are respected.

In the end, it seems that the efforts largely succeeded in this respect. The standard is somewhat difficult to implement (particularly using software-based solutions for general-purpose CPUs), but apparently not excessively so. Chips that can decode video inexpensively in real time at HDTV resolution are now available, software implementations have been demonstrated for CPUs and DSPs, and implementations even exist for embedded hand-held platforms such as mobile phones. Over time, the complexity concerns should further diminish toward relative unimportance.

At the top level of coarseness, estimates of the implementation complexity of decoders for the new standard seem to fall roughly around 3-4 times the complexity of implementation of MPEG-2 video (a roughly ten-year-old standard), or 2-3 times that of H.263 video[22] or the commonly-implemented profiles of MPEG-4 Part 2 video[23] (a roughly five-year-old standard). The complexity ratio for good-quality *encoders* may be somewhat higher, although creative encoder designers can find ways to obviate some of the better-known complex methods of performing encoding processes for the standard. Some studies of complexity of the new standard and its various components have been published[7, 15-19].

Another aspect of the usual understanding of "complexity" concerns how difficult it is to test and debug an implementation, and to understand the design completely and get all of its details completely correct in order to achieve full bug-free interoperability. By such measures, most people would probably consider H.264/MPEG-4 AVC to be a "complex" standard. It contains more than three hundred pages of densely-drafted technical content describing features that can be combined in a highly flexible fashion at the discretion of the encoder. The standard is written in a style intended for expressing precise details in a highly-disciplined fashion – it is not necessarily drafted in a manner that is easy to read or offers adequate tutorial help to non-experts. Implementers are to be cautioned to hire the best experts they can find, pay attention to every detail, question their assumptions, and test their work thoroughly for interoperability with other implementations. (Perhaps it should be noted that the MPEG-2 video standard is also a couple of hundred pages long and the MPEG-4 visual standard is well over five hundred pages, and those are also technically dense.)

4. PATENT LICENSING

Digital video coding technology, being an area of rapid technical innovation and increasingly-high market value over the last 15 years or longer, has been the subject of a significant number of patents. Standards bodies such as ITU-T and ISO/IEC have long-standing policies for dealing with such patent issues, and have decided that patented designs can only be included in their standards when (as far as the developing committee is aware) the patents that are essential for implementation of the standard are available for world-wide licensing on a "reasonable and non-discriminatory" basis. However, the negotiations over the detailed terms of any such licensing activity are kept outside of the committee that develops the standard. It thus falls to the industry as a whole to develop the necessary patent licensing programs.

To a company that wants to implement a design, the burden of understanding what licenses may be necessary and the negotiations for the terms of that licensing can become especially cumbersome if many patents held by a variety of entities apply to the technology (and the consequences of not obtaining a proper license for patents that apply to the company products can be severe). As a result, the concept of *patent pooling* has arisen in recent years, so that a single license can easily be obtained that gives rights to many (ideally, all) of the patents necessary to implement a design. For example, a company named MPEG-LA, LLC. (not to be confused with the MPEG standardization committee) offers pooled licenses for a number of technology designs, including some of those developed by MPEG (hence the naming similarity) and others. Such pools offer *non-exclusive* licenses (i.e., it remains possible to obtain a license from the patent holders without using the pool), and patent holders are not forced to offer licenses through such a pool, but a pool offers an effective "one-stop shop" to simplify the process of licensing the patents in its portfolio. Recently, MPEG-LA, LLC has established such a patent pool for licensing patents it considers to apply to the H.264/MPEG-4 AVC specification[32]. Additionally, Via Licensing Corp. (a division of Dolby Labs) has also announced a pool for additional patents relating to the standard. The licensing issues surrounding the specification may take a while to be completely settled (perhaps about 20 years may be a good estimate), and some dissatisfaction has been expressed with the licensing terms announced by these organizations (primarily centered around aspects sometimes called "usage fees", as contrasted with fees associated with the quantity of deployed encoding and decoding product units). However, the licensing issues do not seem to be a major hindrance to widespread deployment of the standard, as the deployment efforts seem to have developed a strong momentum.

5. DEPLOYMENTS AND ADOPTIONS

This section discusses adoptions and deployments of the H.264/MPEG-4 AVC standard by various application standards bodies, industry consortia, broadcasters, etc. Also discussed are custom integrated-circuit chips and real-time broadcast-quality encoders with high-definition (HD) capability, as the development of such products is an enabling technology for deployment in a number of applications. The information contained in this section has been collected on a best-effort basis from various press releases, press articles, and remarks made and recorded at JVT meetings (with no subsequent indication to the JVT that the remarks are in error after months of meeting report availability). Software product implementations (whether for general-purpose CPUs or more specialized DSPs), non-HD decoder products, and non-real-time, non-broadcast-quality and non-HD encoder products are not included in the discussion, as there is a large number of such products and they are thus more difficult to survey and track.

5.1 Videoconferencing

The new video coding standard was immediately adopted into the relevant ITU-T videoconferencing systems standards (e.g., H.320, H.323, H.324, by way of a new standard called H.241), and the videoconferencing industry began deploying the new standard as soon as it was finished (perhaps sooner). The implementations in the products of the two companies with the largest market share (Polycom and Tandberg) became interoperable around April of 2004, and most other companies with products in that market have also rapidly deployed the new design as well.

5.2 Direct-broadcast satellite television

Based on MPEG-2 Systems multiplexing, much of the direct-broadcast satellite market has announced an intent to embrace the new video coding standard, particularly for HDTV services. Announcements include adoptions by the following:

– BSkyB (UK and Ireland)
– Echostar Dish Network and Voom TV (USA)
– Euro1080 (Europe)
– News Corp. DirecTV (USA)
– Premiere (Germany)

These announcements were made in early 2005.

5.3 Terrestrial-broadcast television

The digital video broadcast (DVB) organization approved the use of the new standard for broadcast television, and the prime minister of France announced the selection of the new standard (exclusively) for all HDTV and subscription-based SDTV digital television services. These events occurred in late 2004. The Advanced Television Systems Committee (ATSC) for U.S. television broadcast is preliminarily adopting the new video standard for its future "robust-mode" back-up channel broadcasting technology.

5.4 Disc storage formats

Both of the major consortia planning to deploy HD next-generation disc formats (the HD-DVD format of the DVD Form and the Blu-ray Disc format of the Blu-ray Disc Association) have announce that support for the new format will be a mandatory player feature in their designs (along with MPEG-2 video and the draft SMPTE standard known as VC-1 and based on Microsoft's Windows Media Video 9). The Sony-designed Universal Media Disc (UMD) video format also specifies the use of H.264/MPEG-4 AVC, and the new Sony PlayStation Portable (PSP) is a well-known new end-user product that includes UMD playback capability and had been shipped in roughly a million-unit deployment by the time of preparation of this manuscript.

5.5 Mobile-device broadcast services

ARIB in Japan and its major broadcasters NHK, Tokyo Broadcasting System (TBS), Nippon Television (NTV), TV Asahi, Fuji TV, and TV Tokyo have announced the intent to use the new standard for mobile-segment video broadcast. The digital multimedia broadcast (DMB) service in the Republic of Korea will also use the new standard.

5.6 Mobile telephony network services

The third-generation partnership project organization (3GPP) adopted the new standard into release 6 of its design. Its counterpart organization 3GPP2 has also made the new standard optional for streaming video services.

5.7 Internet video

The Internet engineering task force (IETF) approved an RTP payload packetization format (RFC 3984) that was specified for the new video coding standard by the IETF Audio-Video Transport (AVT) working group. The Internet streaming media alliance (ISMA) announced the adoption of the new standard into its draft Release 2.0 specification.

5.8 Cross-industry organizations

Two organizations have led a variety of industry efforts toward embracing the standard: the MPEG Industry Forum (MPEGIF, not to be confused with the MPEG standardization body) and the International Multimedia Telecommunications Consortium (IMTC). In this regard, the MPEGIF (formerly known as the MPEG-4 Industry Forum, or M4IF) is perhaps best known for its interoperability testing program, and the IMTC for its generous support in the form of hosting the JVT experts group's email discussion lists and document archives. Both of these organizations were also instrumental in encouraging the formation of reasonable patent licensing programs for the new standard, and the MPEGIF has served a significant promotional marketing role and hosted some relevant technical discussion email groups.

The AVC Alliance, a new organization devoted to promotion and deployment of the new standard, has (more recently) been formed. The digital living network alliance (DLNA), a relatively-new organization for establishing technology compatibility across the consumer electronics, mobile, and personal computing industries, has reportedly included the new standard in its first-version specifications.

5.9 Government use

The United States Department of Defense Motion Imagery Standards Board (MISB) has adopted the new standard as its preferred video coding format for essentially all applications.

5.10 Custom integrated-circuit HDTV decoders

Four companies appear to now (or imminently) be offering custom integrated circuit chips capable of real-time HDTV decoding in real time for the new standard. These products include the following, with some feature highlights as listed:

- Broadcom BCM7411
    - H.264/MPEG-4 AVC Main & High Profile @ Level 4.1
    - HD or dual SD
    - Also decodes MPEG-2
    - And post-processing
- Conexant / Amphion CX2418x
    - H.264/MPEG-4 AVC Main & High Profile @ Level 4
    - Also decodes MPEG-2
    - Also triple-DES encryption/decryption and error/loss concealment
- Sigma Designs SMP8630
    - H.264/MPEG-4 AVC Main Profile @ Level 4
    - Also decodes MPEG-2 (dual) & VC-1
    - Security, audio, and graphics features
- ST Micro STB7100
    - H.264/MPEG-4 AVC Main & High Profile @ Level 4
    - Also decodes MPEG-2 & VC-1
    - Multi-stream capability

5.11 Broadcast-market real-time HDTV encoders

Five companies appear to now (or imminently) be offering real-time broadcast-quality HDTV encoders for the new standard. These products include the following (feature details are not listed, as the degree of public information about these products varies considerably, while trying to maintain a balanced treatment seemed advisable):

- Harmonic / Divicom
- Modulus Video (using some core technology from LSI Logic)
- Scientific Atlanta
- Tandberg Television
- Thomson / Grass Valley

6. CONCLUSIONS

The new video coding standard known as H.264/MPEG-4 Advanced Video Coding (AVC), now in its fourth version, has demonstrated significant achievements in terms of coding efficiency, robustness to a variety of network channels and conditions, and breadth of application. The recent fidelity range extensions have further improved compression quality and further broadened the range of applications, and the recent corrigenda have excised the inevitable errata of the initially-approved versions of the specification. Patent licensing programs have begun, the standard has been adopted into a variety of application specifications, and products suitable for widespread deployment have begun to appear. New work toward the near-term development of scalable video coding (SVC) extensions is also under way.

REFERENCES

1. Joint Video Team (JVT) of ITU-T and ISO/IEC JTC 1, *Advanced Video Coding: ITU-T Rec. H.264* and *ISO/IEC 14496-10*, version 1: May 2003, version 2: March 2004, version 3: July 2004, version 4: January 2005.

2. ITU-T and ISO/IEC JTC 1, *Generic coding of moving pictures and associated audio information – Part 2: Video*: *ITU-T Rec. H.262 and ISO/IEC 13818-2 (MPEG-2)*, Nov. 1994 (with several subsequent amendments and corrigenda).

3. G. J. Sullivan and T. Wiegand, "Video Compression – From Concepts to the H.264/AVC Standard", *Proceedings of the IEEE*, vol. 93, no. 1, January, 2005, pp. 18-31.

4. G. J. Sullivan, P. Topiwala, and A. Luthra, "The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions", *SPIE Annual Conference on Applications of Digital Image Processing XXVII, Special Session on Advances in the New Emerging Standard H.264/AVC*, Aug., 2004, pp. 454-474.

5. T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, number 7, July 2003, pp. 560-576.

6. A. Puri, T. Chen, and A. Luthra, "Video coding using the H.264/MPEG-4 AVC compression standard", *Signal Processing: Image Cummunication*, Elsevier Press, vol. 19, issue 9, October 2004.

7. J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video Coding with H.264/AVC: Tools, Performance, and Complexity", *IEEE Circuits and Systems Magazine*, First Quarter, 2004, pp. 7-28.

8. I. E. G. Richardson / VCodex, "H.264 / MPEG-4 Part 10 Tutorials", available at http://www.vcodex.com/h264.html.

9. I. E. G. Richardson, *H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia*, John Wiley and Sons, September, 2003.

10. S. Okubo, S. Kadono, Y. Kikuchi, and T. Suzuki, (*Eds.*), *et al.*, *H.264/AVC Textbook* (in Japanese), Impress publishers, 2004.

11. B. Erol, A. Dumitraş, F. Kossentini, A. Joch, and G. J. Sullivan, "MPEG-4, H.264/AVC, and MPEG-7: New Standards for the Digital Video Industry", chapter to appear in *Handbook of Image and Video Processing*, 2nd Ed. (A. Bovik, *Ed*.), Academic Press, 2005.

12. M. Flierl and B. Girod, *Video Coding with Superimposed Motion-Compensated Signals: Applications to H.264 and Beyond*, Kluwer, 2004.

13. A. Luthra, G. J. Sullivan, and T. Wiegand (Eds.), "Special issue on the H.264/AVC video coding standard," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, number 7, July 2003.

14. T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-Constrained Coder Control and Comparison of Video Coding Standards", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, number 7, July 2003, pp. 688-703.

15. M. Horowitz; A. Joch; F. Kossentini,; A. Hallapuro, "H.264/AVC baseline profile decoder complexity analysis", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, number 7, July 2003, pp. 704-716.

16. V. Lappalainen, A. Hallapuro; T. D. Hamalainen, "Complexity of optimized H.26L video decoder implementation", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, number 7, July 2003, pp. 717-725.

17. S. Saponara, K. Denolf, G. Lafruit, C. Blanch, and J. Bormans, "Performance and Complexity Co-Evaluation of the Advanced Video Coding Standard for Cost-Effective Multimedia Communications", *EURASIP Journal on Applied Signal Processing*, Issue 2 (2004), pp. 220-235.

18. S. Saponara, C. Blanch, K. Denolf, and J. Bormans, "The JVT Advanced Video Coding Standard: Complexity and Performance Analysis on a Tool by Tool Basis", *Packet Video (PV)* conference, Nantes, 2003.

19. M. Ravasi and M. Mattavelli, "High Abstraction Level Complexity Analysis and Memory Architecture Simulations for Multimedia Algorithms", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 15, No 5, May 2005.

20. S. Wenger, "H.264/AVC over IP", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, number 7, July 2003, pp. 645-656.

21. T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, number 7, July 2003, pp. 657-673.

22. ITU-T, *Video coding for low bit rate communication: ITU-T Rec. H.263*; v1: Nov. 1995, v2: Jan. 1998, v3: Nov. 2000.

23. ISO/IEC JTC 1, *Coding of audio-visual objects – Part 2: Visual: ISO/IEC 14496-2 (MPEG-4 Part 2)*, Jan. 1999 (with several subsequent amendments and corrigenda).

24. Joint Video Team (JVT) of ITU-T and ISO/IEC JTC 1, *JVT Document Archive Site*, http://ftp3.itu.int/av-arch/jvt-site.

25. Video Coding Experts Group (VCEG) of ITU-T Study Group 16 (Question 6), *VCEG Document Archive Site*, http://ftp3.itu.int/av-arch/video-site.

26. K. Sühring (Ed.) and the Joint Video Team (JVT) of ITU-T and ISO/IEC JTC 1, *H.264/AVC Joint Model (JM) Software Coordination Site*, http://iphome.hhi.de/suehring/tml/.

27. ITU-R, *Methodology for the Subjective Assessment of the Quality of Television Pictures: ITU-R Rec. BT.500-11*, June, 2002.

28. ISO/IEC JTC 1/SC 29/WG 11 (MPEG), "Report of The Formal Verification Tests on AVC (ISO/IEC 14496-10 | ITU-T Rec. H.264)", MPEG document number N6231 (available at http://www.chiariglione.org/mpeg/quality_tests.htm), Waikoloa, Hawaii, December, 2003.

29. T. Oelbaum, V. Baroncini, T. K. Tan, and C. Fenimore, "Subjective Quality Assessment of the Emerging AVC/H.264 Video Coding Standard", *International Broadcast Conference (IBC)*, Amsterdam, Netherlands, 2004.

30. C. Fenimore, V. Baroncini, T. Oelbaum, and T. K. Tan, "Subjective Testing Methodology in MPEG Video Verification", *SPIE Annual Conference on Applications of Digital Image Processing XXVII, Special Session on Advances in the New Emerging Standard H.264/AVC*, Aug., 2004, pp. 503-511.

31. T. Wedi and Y. Kashiwagi, "Subjective quality evaluation of H.264/AVC FRExt for HD movie content", Joint Video Team document JVT-L033, July, 2004.

32. L. A. Horn, "AVC/H.264 Patent Portfolio License", *SPIE Annual Conference on Applications of Digital Image Processing XXVII, Special Session on Advances in the New Emerging Standard H.264/AVC*, Aug., 2004, pp. 495-502.