NEW STANDARDIZED EXTENSIONS OF MPEG4-AVC/H.264 FOR PROFESSIONAL-QUALITY VIDEO APPLICATIONS

Gary J. Sullivan [1], Haoping Yu [2], Shun-ichi Sekiguchi [3], Huifang Sun [3], Thomas Wedi [4], Steffen Wittmann [4], Yung-Lyul Lee [5], Andrew Segall [6], and Teruhiko Suzuki [7]

[1] Microsoft Corporation, garysull@microsoft.com
[2] Thomson Inc., haoping.yu@thomson.net
[3] Mitsubishi Electric Corporation, sekiguchi.shunichi@eb.mitsubishielectric.co.jp, hsun@merl.com
[4] Panasonic R&D Center Germany, {thomas.wedi, steffen.wittmann}@eu.panasonic.com
[5] Sejong University, yllee@sejong.ac.kr
[6] Sharp Labs of America, Inc., asegall@sharplabs.com
[7] Sony Corp, teruhiko@av.crl.sony.co.jp

ABSTRACT

To support high quality video applications, the Joint Video Team (JVT) has recently added five new profiles, two new supplemental enhancement information (SEI) messages, and two new extended gamut color space indicators to the MPEG4-AVC/H.264 video coding standard. The new profiles include substantial feature enhancements for high-quality video applications, including improved-efficiency 4:4:4 video format coding, improved-efficiency lossless macroblock coding, coding 4:4:4 video pictures using three separately-coded color planes, and support of bit depths up to 14 bits per sample. The new features were developed to support a wide range of applications where high quality video compression is demanded, including professional and semi-professional scenarios in particular. They also anticipate the introduction of higher fidelity displays. In this paper, the new extensions are presented along with quantitaive estimates of the benefits of the new features and a discussion of the target application environments.

Index Terms— *MPEG-4, AVC, H.264, 4:4:4 video coding, Intra-only video coding, professional video applications.*

1. INTRODUCTION

To respond to strong demands for compression technology appropriate to high quality applications such as studio camcorders, professional digital video recording and editing systems, digital cinema / large-screen digital imagery, high fidelity display systems, etc., the Joint Video Team (JVT) has recently developed two new amendments to the MPEG4-AVC/H.264 standard that specify a set of five new profiles, two new supplemental enhancement information (SEI) messages, and two new extended gamut color space indicators.

The new profiles are called the High 10 Intra, High 4:2:2 Intra, High 4:4:4 Intra, CAVLC 4:4:4 Intra and High 4:4:4 Predictive profiles. They include substantial feature enhancements for high-quality video applications, including improved-efficiency 4:4:4 video format coding, improved-efficiency lossless macroblock coding, coding 4:4:4 video pictures using three separately-coded color planes, and bit depths up to 14 bits per sample.

In addition, all-Intra coding (i.e., the coding of each individual picture separately without prediction from the content of other pictures) is widely used in content creation, production, and post-production applications, primarily for ease of editing. Four of the five new profiles have thus been defined for all-Intra coding to enable applications requiring simple editing of bitstreams without imposing the burden of implementing the extra inter-picture decoding functionality that would not be needed in such applications.

To further enhance the decoded image quality on displays even with different characteristics, new "post-filter hint" and "tone mapping" SEI messages are also defined in the new extensions. Additionally, to support future high-quality display technology, the ability to indicate the use of the extended gamut color spaces defined by IEC 61966-2-4 ($xvYCC_{601}$ and $xvYCC_{709}$) and ITU-R BT.1361 has been enabled.

2. TECHNICAL DESIGNOF THE EXTENSIONS

2.1 Enhanced 4:4:4 compression capability

Recently, the 4:4:4 video sampling format (i.e., the use of the same picture sampling resolution for all three components of the sample arrays used to represent video brightness and color information) has been getting increasing attention in many high quality video applications. Due to the nature of the 4:4:4 format, specialized compression techniques have proved beneficial to reduce the data rate to meet the bandwidth and storage capacity requirements imposed by these demanding applications.

The design of the new High 4:4:4 Intra, CAVLC 4:4:4 Intra, and High 4:4:4 Predictive profiles was developed based on the determination that overall compression performance in 4:4:4 coding could be improved significantly by applying the type of signal processing that had previously been used only for the luma component in the original MPEG4-AVC/H.264 standard to all three of the color components (e.g., the red, green and blue, or Y, Cb, and Cr components) that are used to represent 4:4:4 visual information. In the other profiles of MPEG4-AVC/H.264, one color component (designated as the luma component), is coded using more powerful signal processing techniques, while the other two components (designated as the chroma components) are processed using less sophisticated techniques that were developed originally for the consumer-grade 4:2:0 video format (in which the chroma components are subsampled by a factor of two both horizontally and vertically). When the input to the codec is in the 4:4:4 format, the quadrupling of the number of non-primary samples resulted in spending more bits in those two channels and brought to light the signal processing deficiencies that would not matter in other sampling scenarios. This particular problem is more noticeable in the Intra coded frames since the spatial prediction method in the MPEG4-AVC/H.264 chroma coding toolbox is far less powerful than that in the coding toolbox for luma. According to various test result reports, these improved coding techniques can reduce the bit rate needed for high fidelity video coding (e.g., for PSNR at or above 45 dB) by around an average of 20%.

The design of the new 4:4:4 profiles, as specified in the "New profiles for professional applications" amendment to ITU-T Rec. H.264 & ISO/IEC 14496-10 (Amendment 2 to the 2005 edition [1]), supports any of the color spaces listed in the standard, i.e., RGB, YCbCr or YCgCo with or without extended gamut support, and has two basic operation modes depending upon the value of a new flag added to the bitstream called the separate_colour_plane_flag. When this flag is equal to 0, the samples from the three color components in a macroblock are encoded together using shared slice structures, prediction mode selections and motion representations. In contrast, when separate_colour_plane_flag is equal to 1, each color component array is treated as a separate monochrome video picture and is encoded individually using its own slice segmentation and mode selections. These two 4:4:4 coding modes are thus referred to as common mode (separate_colour_plane_flag = 0) and independent mode (separate_colour_plane_flag = 1) in this paper. Since both basic 4:4:4 operational modes use the general coding structure and existing (luma) coding tools of the prior design, the extra implementation effort to support these profiles in an implementation supporting the MPEG4-AVC/H.264 non-4:4:4 profiles is minimized. In various performance tests, the compression capability of the common and independent modes has been reported to be about the same (at maximum encoder implementation complexity) when coding video with high fidelity.

A new lossless DPCM macroblock type is also added for intra-picture coded residues under both operation modes in the three new profiles that support 4:4:4 video. In addition, the bit depth capability of these three profiles is extended up to 14 bits per sample. In these profiles, the lossless DPCM and enhanced bit depth support can also be used with other chroma formats (i.e., 4:2:2, 4:2:0, and monochrome).

In the following discussion, the common mode, the independent mode, and the new lossless DPCM process are discussed separately. The details of the new 4:4:4 profiles are specified in [1].

2.1.1 Common and independent coding modes

The design principle used in the development of the new 4:4:4 profile was to achieve maximum coding efficiency while maximally using the existing coding architecture and coding tools.

There are two operational coding modes in the new 4:4:4 profiles. One is the common mode, where a macroblock consists of samples from all three color components and the bitstream content is characterized by a single slice structure and a single set of macroblock-level coding parameters, such as macroblock types, prediction modes, motion vectors, and so on. In the common mode, no change is made to the macroblock types and a common prediction mode is used for all color components in a macroblock in order to minimize the necessary syntax changes to the existing specification and to reduce the algorithm complexity. Residual signals of three color components obtained by intra/inter prediction are coded with entropy coding in a manner analogous to what was previously used only for the luma component.

The other coding mode is referred to as the independent mode, which re-uses the existing MPEG4-AVC/H.264 monochrome coding algorithm to encode each color channel of 4:4:4 source video independently as a separate picture [2], [3]. To accommodate the powerful coding tools of MPEG4-AVC/H.264 such as inter prediction or CABAC for new applications requiring super high-resolution video, the independent mode provides an enhanced parallel processing solution for the encoder by allowing a splitting of the entire encoding processes of the three color channels. In addition, the independent mode can capture local statistics of each color channel well, at the expense of an increased bit budget for the channel-specific side information. It can be particularly useful for encoders in those applications that require high-fidelity coding (where the extra bits needed for the side information matter less).

2.1.2 Lossless DPCM process for Intra-coded residues

To improve the compression efficiency for the lossless case, a new Intra residual DPCM process is included in both the common mode and the independent mode. This process is applied to the intra coded macroblocks having a spatial prediction mode that is either horizontal or vertical. In other cases, the efficient lossless macroblock representation capability is fulfilled by a technique known as transform-bypass lossless coding [4]. Further information about the new DPCM process can be found in [1], [5] and [6].

2.2 New profiles

A total of five new profiles have been added in this new amendment, as diagrammed below in Fig. 1 along with two of the existing profiles (dotted-line region) for context relationship. To reduce decoder complexity, all of the new profiles facilitate decoder parallel-processing by imposing a constraint that the maximum size of a slice cannot exceed one fourth of the maximum supported picture size.

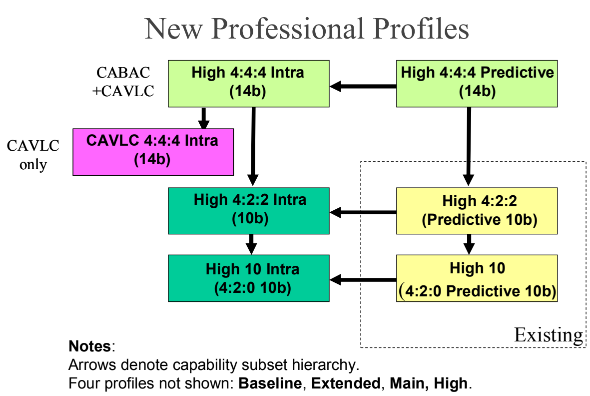

Fig. 1: Codec using post-filter hint SEI message

2.2.1 *Intra profiles*

For professional applications, all-intra coding is a common coding method, primarily for ease of bitstream editing. Furthermore, several experiments and demonstrations have shown that the intra coding design of MPEG4-AVC/H.264 has an excellent performance, even when compared to the state of the art in still image coding schemes. Thus, the new extensions include four profiles for all-intra applications. Each of these is designed as an all-intra subset of a corresponding predictive profile, and an additional profile is defined to reduce decoder complexity by restricting the entropy coding method only to the less complex of the two entropy coding schemes (i.e., to CAVLC).

2.2.2 *High 4:4:4 Predictive profile*

As the superset of the existing High 10 and High 4:2:2 profiles, the new High 4:4:4 Predictive profile, which supports all of the coding tools (including those used in the inter frame prediction) and bit depths up to 14 bits per sample, has been created so that real-time transmission or efficient storage of high-quality video can also be achieved. As the with new High 4:4:4 Intra profile, the High 4:4:4 Predictive profile also supports the common and independent modes and is applicable to wider variety of 4:4:4 video applications.

2.3 New SEI messages

2.3.1 *Post-Filter Hint SEI Message*

In the field of video coding, it is common to apply post-filtering in order to enhance the quality of decoded video signals. An optimal post-filter can be designed by minimizing the error between the original uncompressed source signal and the post-filter output at the decoder. This approach requires a-priori information of the statistics of the original and the decoded signal that is available only to the encoder.

In order to provide this necessary information to a decoder, the so-called post-filter hint SEI message is included in the new MPEG4-AVC/H.264 extensions. With this SEI message it is possible to transmit the coefficients of a post-filter calculated at the encode side or to transmit the correlation information between the original and the decoded signal, which is necessary to design a optimal post-filter (Wiener filter) on decoders. Fig. 2 shows the structure of a codec that applies this concept [7].

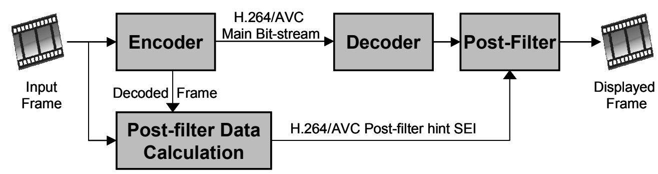

Fig. 2: Codec using post-filter hint SEI message

2.3.2 *Tone mapping SEI message*

A final component of the new amendment is the tone mapping SEI message. This message supports the use of the higher bit-depth profiles for storing image sequences with a large luminance dynamic range. This is in contrast to the more traditional approach of utilizing higher bit-depths to increase the numerical dynamic range of the representation. Supporting these alternative applications requires no changes in the coding process. Instead, the only difficulty is that a display device needs additional information on how to render the decoded, higher bit-depth sequence.

The tone mapping SEI message provides the necessary information to a display device [8]. This is accomplished by transmitting a “tone mapping curve” from the encoder to the display, where the tone mapping operation describes a method for mapping the higher bit-depth sequence to a lower bit-depth display. Several methods for representing the tone map are provided to make the message flexible. These methods include a linear, sigmoidal and piece-wise constant mapping. Additionally, a point-wise lookup table is supported.

3. SIMULATION AND TEST RESULTS

3.1 New High 4:4:4 profiles

In this paper, a sample test result is shown in Figure 2 comparing the new design for 4:4:4 coding vs. a straightforward extension of MPEG4-AVC to the 4:4:4

format without including the enhanced chroma handling capability (referred to as "4:4:4 SE" in the figure) for intra-only and predictive coding scenarios. More test results can be found in [2], [3], and [9] . The source sequence used in the test was a clip from the "Treasure Planet" film scan sequence, and it is in the RGB color space with 10-bits per color component. Since the uncompressed source video is provided in the RGB domain, average R, G, and B PSNR in the RGB domain is used in Fig. 3 to measure the distortion. The coding was done in the RGB color space. The figure clearly show the substantial performance benefit of the new 4:4:4 coding solution

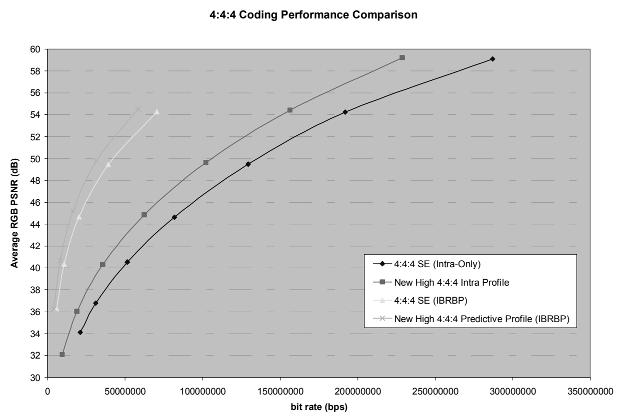

Fig. 3: 4:4:4 coding test results

3.2 Post-filter hint SEI message

Mean coding efficiency improvements of about 0.4 dB, which corresponds to bit rate reductions of about 8.5%, have been observed with this technique compared to MPEG4-AVC/H.264 coding without post-filtering. Fig. 4 shows the bit rate reductions for some 4:4:4 video sequences. More detailed results can be found in [7].

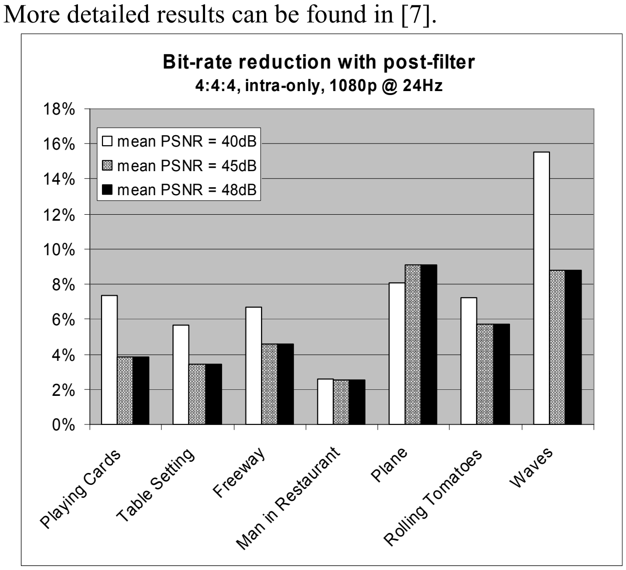

Fig. 4: Bit rate reductions achieved with post-filter hint SEI

4. CONCLUSION

This paper describes the new extensions that have recently been standardized by the JVT for MPEG4-AVC/H.264. The extensions include five new profiles, two new SEI messages, and two extended gamut color space indicators. These extensions were designed for high fidelity video applications. Among the five new profiles, three of them include a new coding design to compress 4:4:4 video. In addition, High 10 Intra and High 4:2:2 Intra profiles are also added to the standard to provide state-of-the-art compression technology to other professional video applications. The new SEI messages can further improve the decoded image quality on displays with various display characteristics. Finally, extended gamut color space indicators address the needs of future enhanced display technologies.

5. REFERENCES

(Note: JVT documents listed below are publicly available on line at http://ftp3.itu.int/av-arch/jvt-site.)

[1] G. J. Sullivan and H. Yu (Editors), "New profiles for professionial applications" JVT document JVT-V204, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Marrakech, Morocco, January 2007.

[2] S. Sekiguchi, Y. Isu, Y. Yamada,K. Asai, and T. Murakami, "Results of Core Experiment on 4:4:4 coding (CE5)", JVT document JVT-S014, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Geneva, March 2006.

[3] S. Sekiguchi, Y. Isu, K. Sugimoto, Y. Yamada, K. Asai, and T. Murakami, On Separate Color-Plane Prediction for Direct 4:4:4 Video Coding", *IEEE Intl. Conf. on Image Proc.* (*ICIP*), TP-L3.8, October 2006.

[4] S. Sun, "New Results of Lossless Coding", JVT document JVT-D028, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Fairfax, Virginia, May 2002.

[5] Y.-L. Lee, K.-H. Han, and G. J. Sullivan, "Improved Lossless Intra Coding for H.264/MPEG-4 AVC", IEEE Trans. Image Proc., vol. 15, no. 9, pp. 2610-2615, Sept. 2006.

[6] Y.-L. Lee and K.-H. Han, "Complexity of the proposed lossless intra for 4:4:4," JVT document JVT-Q035, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Nice, France, October 2005.

[7] S. Wittmann and T. Wedi, "SEI message on post-filter hints", JVT document JVT-U035, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Hanghzou, China, October 2006.

[8] A. Segall, L. Kerofsky and S. Lei, "Tone Mapping SEI Message", JVT document JVT-T060, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Klagenfurt, Austria, July 2006.

[9] H. Yu and L. Liu, "Advanced 4:4:4 Profile for MPEG4-Part10/H.264", JVT document JVT-P017, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Poznan, Poland, July 2005.