HEVC vs. H.264/AVC Standard Approach to Coder’s Performance Evaluation

ZORAN MILICEVIC(1), ZORAN BOJKOVIC(2), KAMISETTY R. RAO(3)

1Department of Telecommunication and IT GS of the SAF, 11000 Belgrade, REPUBLIC OF SERBIA
2University of Belgrade, Studentski Trg 1, 11000 Belgrade, REPUBLIC OF SERBIA
3The University of Texas at Arlington, TX, UNITED STATES of AMERICA

mmilicko@eunet.rs, z.bojkovic@yahoo.com, rao@uta.edu

Abstract: - The H.264/AVC standard achieves much higher coding efficiency than the H.263, MPEG-2 and MPEG-4 standard, due to its improved inter and intra-prediction modes at the expense of higher computation complexity. Throughout the evolution of video coding standards, continued efforts have been made to maximize compression capability and improve other characteristics such as data loss robustness, while considering the practical computational resources. On the other hand, High Efficient Video Coding (HEVC) standard can provide a significant amount of increased coding efficiency compared to previous H.264/AVC standard. The features of the new design provide approximately a 50% bit-rate savings for equivalent perceptual quality relative to the performance of prior standard (especially for a high-resolution video). In order to compare the performance and complexity without significant rate-distortion performance degradation, the two different HEVC coders vs. H.264/AVC coder are tested for the fixed Quantization Parameter (QP) value, when Main profile, appropriate motion vector (MV) search ranges and IPPP structures are used. Simulation results have shown that the bit-rate was reduced over 50%, while the encoding time saving is slightly decreased up to 16% depending on the tested video sequence, when reference HEVC software’s HM-14.0 and HM-15.0 are compared to reference H.264/AVC software JM 18.6. However, there was negligible loss in term Signal-to-Noise Ratio (SNR).

Key-Words: - H.264/AVC standard, HEVC standard, Encoding time saving, Signal-to-noise ratio, Bit-rate reduction

1 Introduction and Motivation

AVC is the dominant video coding design of today. The core design of the standard has given it exceptionally strong compression capability relative to prior designs together with the flexibility and robustness to enable its use in an extremely broad variety of network and application environments. Its several generations of extensions that have been added to its design have further broadened its capabilities to apply in professional-domain environments and have given it highly-flexible scalability features and 3D stereoscopic and multiview support – while retaining a consistency of design approach that make it straightforward to deploy AVC products that support a broad range of applications.

On the other hand, High Efficiency Video Coding (HEVC) increased compression efficiency compared to AVC, with a focus on video sequences with resolutions of HDTV and beyond. In addition to broadcasting applications, HEVC caters towards the mobile market. HEVC provides more flexibility in terms of larger block sizes, more efficient motion compensation and motion vector prediction as well as more efficient entropy coding. With the standards HEVC and 3DV, MPEG and JCT-VC provided codecs to deliver highest quality video content in 2D and 3D. Due to the limitation of bandwidth and stereo TV, markets for the new standards will be developed very soon.

An increasing diversity of services, the growing popularity of High Definition (HD) video, and the emergence of beyond HD formats (e.g., 4kx2k or 8kx4k resolution) are creating stronger needs for coding efficiency superior to H.264/MPEG-4 AVC’s capabilities [1].

The need is even stronger when higher resolution is accompanied by stereo or multiview capture and display. An increased desire for higher quality and resolutions is also arising in mobile applications.

In HEVC, the main goal is to achieve a compression gain as much higher as possible, when compared to the H.264/AVC at the same video quality. To facilitate the functionality, a system that contains hybrid frame buffer compression, low resolution intra prediction, cascaded motion
compensation and in-loop deblocking components has to be developed. On the other hand, a low-resolution decoding framework in the context of the emerging HEVC video coding is introduced. When HEVC decoder is operating in a low resolution mode, the decoding energy is saved. This is necessary for battery powered mobile handsets and other power devices which provide direct benefit to the consumer.

In the family of video coding standards, HEVC has the potential to replace/supplement all the existing standards. While the complexity of the HEVC encoder is several times that of the H.264/AVC, the decoder complexity is within the range of the latter [2]. The implementation complexity of HEVC overall is not a major burden relative to H.264/AVC using a modern processing technology. Several tests have shown that HEVC provides improved compression efficiency up to 50% bit-rate reduction for the same subjective video quality compared to H.264/AVC [3]. This provides direct benefit to the consumer of different multimedia applications.

In this work, our goal is to provide comparison of the performance and complexity without significant rate-distortion performance degradation, when the HEVC vs. H.264/AVC coders are tested for the fixed quantization parameter (QP) value. In our simulation, we have proposed exhaustive tests for both codecs for appropriate parameter sets.

This paper is organized as follows. After an introduction and motivation section II describes H.264/AVC standard background. Section III briefly presents HEVC project and framework, while Section IV contains experimental results and discussion. Section V provides closing remarks.

2 H.264 Standard Background
The Advanced Video Coding (AVC) standard has achieved a significant improvement in compression capability compared to prior standards, and it provides a network-friendly representation of video that addresses both non-conversational (storage, broadcast, or streaming) and conversational (videotelephony) applications. Extensions of AVC have given it efficient support for additional functionality such as scalability at the bitstream level and 3D stereo/multiview coding [4, 5].

The AVC standard was jointly developed by MPEG (the ISO/IEC Moving Picture Experts Group) and the ITU-T Video Coding Experts Group (VCEG). It is published both as ISO/IEC International Standard 14496-10 (informally known as MPEG-4 Part 10) and ITU-T Recommendation H.264 [6, 7]. H.264 fulfills significant coding efficiency, simple syntax specifications, and seamless integration of video coding into all current protocols and multiplex architectures.

H.264 video coding standard has the same basic functional elements as previous standards (MPEG-1, MPEG-2, MPEG-4 part 2, H.261, H.263), i.e., transform for reduction of spatial correlation, quantization for bitrate control, motion compensated prediction for reduction of temporal correlation, entropy encoding for reduction of statistical correlation. However, in order to fulfill better coding performance, the important changes in H.264 occur in the details of each functional element. Some functionalities were introduced in H.264/AVC as: intra-prediction in spatial domain, hierarchical transform with (4x4, 8x8) integer DCT transforms and (2 times 2, 4x4) Hadamard transforms, multiple reference pictures in inter-prediction, generalized bidirectional prediction (forward/forward, backward/backward), weighted prediction, deblocking filter, Context – Based Adaptive Variable Length Coding (CAVLC) entropy coding, parameter setting, flexible macroblock ordering, redundant slices, and SP (Switched P)/SI (Switched I) slices for error resilience.

To address the need for flexibility and customizability, the AVC design covers a video coding layer (VCL), which is designed to efficiently represent the video content, and a network abstraction layer (NAL), which formats the VCL representation of the video and provides header information in a way that enables the coded video to be conveyed by a variety of transport layers or storage media [5, 8].

Technically, the design of the H.264/MPEG4-AVC video coding layer is based on the traditional hybrid concept of block-based motion-compensated prediction (MCP) and transform coding. The encoder runs the same prediction loop as the decoder to generate the same prediction signal and subtracts it from the original picture to generate the residual. The most difficult task of the encoder is to determine its choices for prediction (such as motion vectors, block sizes, and inter/intra prediction modes) and residual difference coding, for which the encoder tries to use as few bits as possible after entropy coding to obtain adequate decoded quality [5].

The improvement in coding performance comes mainly from the prediction part. Intra prediction significantly improves the coding performance of H.264/AVC intra frame coder [9]. On the other side,
inter prediction is enhanced by motion estimation with quarter-pixel accuracy, variable block sizes, multiple reference frames and improved spatial/temporal direct mode [10].

H.264/AVC defines a set of Profiles, each supporting a particular set of coding functions and each specifying what is required of an encoder or decoder that complies with the Profile. Also, profiles are defined to cover the various applications from the wireless networks to digital cinema.

In 2004 Joint Video Team (JVT) added new extensions known as the Fidelity Range Extensions (FRExt), which provide a number of enhanced capabilities relative to the base specification. Also, the Scalable H.264/AVC extension is applied to extend the hybrid video coding approach of H.264/AVC in a way that a wide range of spatial-temporal and quality scalability is achieved. The SVC approach of AVC is based on the “layered coding” principle, which encodes differential information between layers of different quality or resolution [5]. The next major feature added to the standard was Multiview Video Coding (MVC). In its basic design concepts, MVC is an extension of the inter-view prediction principle similar as implemented in the MPEG-2 multiview profile (assuming two or multiple cameras shooting the same scene). Compared to MPEG-2, MVC benefits from the more flexible multiple reference picture management capabilities in AVC [5]. Important applications of AVC are: mobile telephony, mobile TV, mobile video players, TV broadcast, camcorders, Blu-ray disc, videotelephony/videoconferencing, Internet video [11, 12].

3 HEVC Project and Framework

H.264/MPEG-4 AVC is the video coding standard directly preceding the HEVC Project, which has displaced the older standards in its application domain [1]. To assist the interested industry, the standardization effort includes, besides the development of a text specification documents, reference software source code as an example of how HEVC video can be encoded and decoded. A standard test data has to be developed for testing conformance to the standard.

HEVC standard is joint video project of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) standardization organizations, working together in a partnership known as the Joint Collaborative Team on Video Coding (JCT-VC).

At the core, The HEVC standard is designed to achieve multiple goals, such as increased video resolution, coding efficiency, transport system integration and data loss resilience implementability using parallel processing architectures [1]. Corea [13], has carried out a detailed investigation of the performance of coding efficiency versus computational complexity of HEVC encoders. The investigation focuses on identifying the tools that most affect vital parameters: efficiency and complexity.

The HEVC design follows the block-based hybrid video coding approach [14]. Namely, the basic source-coding algorithm is a hybrid of inter-picture prediction to exploit temporal statistical dependencies, intra-picture prediction to exploit spatial statistical dependencies, and transform coding of the prediction residual signals to further exploit spatial dependencies.

There is no signal element in the HEVC design that provides the majority of its significant improvement in compression efficiency [3]. A plurality of smaller improvements adds up to significant gain.

In an encoding algorithm, each picture is split into block-shaped regions, with block positioning being conveyed to the decoder. The first picture of a video sequence is coded using only intra-picture prediction. This type uses some prediction of data spatially from region-to-region within the same picture, but has no dependence on other pictures. For all remaining pictures of a sequence or between random access points, inter-picture temporally-predictive coding modes are used for most blocks.

The encoding process for inter picture prediction consists of choosing motion data comprising the selected reference picture and motion vector (MV) to be applied for predicting the samples of each block. The encoder and decoder generate identical inter prediction signals by applying motion compensation (MC) using the MV and mode decision data which are transmitted as side information. The residual signal of the intra or inter prediction, which is the difference between the original block and its prediction, is transformed by a linear spatial transform. The transform coefficients are then scaled, quantized, entropy coded, and transmitted together with the prediction information. The quantized transform coefficients are constructed by inverse scaling and are then inverted transformed to duplicate the decoded approximation of the residual signal.

The residual is then added to the prediction and the result of that addition may be then fed into one or two loop filters to smooth out artifacts induced by the block-wise processing and quantization. The final picture representation is stored in a decoded
picture buffer to be used for the prediction of subsequent picture.

The analogous structure in HEVC is the coding tree unit (CTU). It has a size selected by the encoder and can be larger than traditional macroblock. The CTU consists of a luma coding tree block (CTB) and the corresponding chroma CTBs and syntax elements [15]. For each prediction block (PB), either one or two motion vectors can be transmitted resulting either in uni-predictive or bi-predictive coding, respectively.

The decoded boundary samples of adjacent blocks are used as reference data for spatial prediction in PB regions, when inter-picture prediction is not performed. The selected intra-prediction modes are encoded by deriving most probable modes based on those of previously decoded neighboring PBs.

Uniform reconstructed quantization is used with quantization scaling matrices supported for the various transform block sizes. Context Adaptive Binary Arithmetic Coding (CABAC) is used for entropy coding. Comparing to the CABAC scheme in H.264/AVC, several improvements exist subject to throughput speed, compression performance, and reducing context memory requirements.

In loop deblocking filtering (DF), the design of DF is simplified in regard to its decision-making and filtering processes. At the same time, it is made more friendly to parallel processing. After the deblocking filter, a non-linear mapping is introduced in the inter-picture prediction loop. When using sample adaptive offset (SAO), the goal is to better reconstruct the original signal amplitudes.

HEVC address essentially all existing applications of H.264/MPEG-4 AVC and has been designed to increased video resolution and increased use of parallel processing architectures [16, 17].

4 Experimental Results and Discussion
To evaluate the performance of the H.264/AVC and HEVC codecs, H.264/AVC reference software JM 18.6 [18] and HEVC reference software HM-14.0, as well as, reference software HM-15.0 [19] are tested with fixed Quantization Parameter value (QP). The system platform is the Intel(R) Core(TM) i3-2328M Processor of speed 2.2 GHz, 6 GB RAM, and Microsoft Windows 7 Professional. The configuration of H.264/AVC was as follows: (1) Main profile, (2) four values of Levels: 2.1, 3.1, 4.0 and 5.0 (3) Quantization Parameters (QP) value is 28, (4) MV search range is 16, (5) Period of I-pictures: only first, (6) R-D optimization is selected like high complexity mod (7) Reference frame number equals to 4, (8) Context Adaptive Binary Arithmetic Coding (CABAC) is enabled, (9) Hadamard transform is used (10) Group of picture structure is IPPP, (11) The number of frames in a sequence is 100 and (12) Types of pictures are I and P. On the other hand, the HEVC configurations were as follows: (1) Main profile, (2) four values of Levels: 2.1, 3.1, 4.0 and 5.0, (3) P pictures, (4) period of I-pictures: only first, (5) Hadamard transform was used, (6) MV search range was 64, (7) SAO, AMP and RDOQ were enabled, (8) GOP length 4 in IPPP format was used. The QP used was 28.

Comparisons with the case of exhaustive search were performed with respect to the change of average signal to noise ratio - SNR (ASNR), the change of average data bits (ABit-rate), and the change of average encoding time (ΔTime), respectively.

In order to evaluate the timesaving of the both algorithms, the following calculation is defined to find the time differences. Let $T_{H.264}$ denotes the coding time used by JM18.6 encoder and $T_{HEVC}$ be the time taken by the HM-14.0. The time difference is defined as:

$$\Delta Time = \frac{T_{HEVC} - T_{H.264}}{T_{H.264}} \times 100\%$$

A group of experiments were carried out on the recommended sequences with quantization parameter QP=28. We chose QP=28 as value of the QP, because it is approximately average value in reference software’s. The average Δ bit-rate is the bit rate difference expressed as a percentage between JM 18.6 encoder and the HM-14.0, as well as, the HM-15.0, respectively. The SNR values of luma (Y) component of pictures are used. The average Δ SNR is the SNR difference expressed as a percentage between JM 18.6 encoder and the HM-14.0, as well as, the HM-15.0, respectively.

The selected test sequences are in Standard Definition (SD), High Definition (HD) and Full High Definition (Full HD). The test sequences have been selected to emphasize different kind of motions and contents. For the experiments, we used the first 100 frames of the 4 different test sequences in different recommended classes (Basketball Pass (416x240 pixels) - class D, Party Scene (832x480 pixels) – class C, City (1280x720 pixels) – class E and Kimono (1920x1080 pixels) – class B) [1].

Starting from the fact that video coding standard HEVC hold CABAC entropy coding method, we have used CABAC [8] in H.264/AVC because this is entropy coding tool for Main profile.
We measured SNR only for Y because human visual system is more sensitive to luma then to chroma components of pictures. Also, we have applied Hadamard transformation because it improves the encoder performance comparing to other transformations [20].

Table 1 shows the performance of the compared reference codecs for P pictures processing in the IPPP structure for QP=28, respectively, based on our simulation results, which have shown the efficiency of the approach in order to analyze both reference codecs performance.

When various test sequences in different formats are processed, the encoding time saving and bit-rate are reduced, while there is negligible loss in term SNR for luma component of picture by HEVC codec. The bit-rate is reduced more than 50%, while the encoding time saving is slightly decreased in average -8.38% (HD and full HD test sequences are reduced over 14% and 11%, respectively) when HM-14.0 is compared to reference software JM 18.6.

Table 1. Experimental results for P pictures in the IPPP format and QP=28 when HEVC (HM-14.0) and H.264/AVC (JM-18.6) are compared.

<table>
<thead>
<tr>
<th>Test sequences</th>
<th>Format</th>
<th>Δ Time (%)</th>
<th>Δ SNR - Y (dB)</th>
<th>Δ Bit rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basketball Pass</td>
<td>SD (416x240)</td>
<td>-8.19</td>
<td>3.19</td>
<td>-48.98</td>
</tr>
<tr>
<td>Party Scene</td>
<td>SD (832x480)</td>
<td>-2.90</td>
<td>5.55</td>
<td>-50.21</td>
</tr>
<tr>
<td>City</td>
<td>HD (1280x720)</td>
<td>-15.94</td>
<td>1.88</td>
<td>-58.45</td>
</tr>
<tr>
<td>Kimono</td>
<td>Full HD</td>
<td>-6.16</td>
<td>2.06</td>
<td>53.82</td>
</tr>
<tr>
<td></td>
<td>(1920x1080)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>-8.28</td>
<td>3.17</td>
<td>-54.54</td>
</tr>
</tbody>
</table>

When experimental results shown in Table 1 and Table 2 are compared, it can be concluded that performance HEVC HM-14.0 and HM-15.0 encoder’s vs JM-18.6 encoder provide very similar output results in term of SNR, bit-rate reduction and encoding time saving.

In Fig. 1 (A) SNR curves are depicted for Basketball Pass test sequence in 416x240 resolutions, in which the SNR-YUV is plotted as a function of the frame number for both tested encoders. Also, In Fig. 1 (B), curves are depicted for Party Scene test sequence in SD resolution. Next, in Fig. 1 (C), curves are depicted for City test sequence in HD resolution. Finally, in Fig. 1 (D), curves are depicted for Kimono test sequence in full HD resolution. These results indicate that there is negligible loss in term SNR for luma component of picture when HM-14.0 and HM-15.0 encoder is compared with JM 18.6 encoder.

On the other hand, Table 2 shows the performance of the compared reference codecs JM-18.6 and HM-15.0 for the same condition as in previous test case. In this case, there is also bit-rate reduction more than 50%, while the encoding time saving is slightly decreased in average -8.28. If we focus on SNR, it is obvious that there is also negligible loss for luma component of picture by HEVC codec.

Table 2. Experimental results for P pictures in the IPPP format and QP=28 when HEVC (HM-15.0) and H.264/AVC (JM-18.6) are compared.
Fig. 1. SNR curves when HEVC (HM-14.0 and HM-15.0) is compared with H.264/AVC (JM-18.6) for Basketball Pass (A), Party Scene (B), City (C) and Kimono1 (D) test sequences.

However, in Fig. 2 bit-rate savings curves are depicted for four typical tested sequences. These results indicate that the emerging HEVC (HM-14.0 and HM-15.0) standard encoder clearly outperforms its predecessors in terms of coding efficiency (it provides over a 50% bit-rate savings for negligible degradation perceptual quality) for all tested applications.
Fig. 2. Bit-rate curves when HEVC (HM-14.0 and HM-15.0) is compared with H.264/AVC (JM-18.6) for Basketball Pass (A), Party Scene (B), City (C) and Kimono1 (D) test sequences.

Fig. 3 and Fig. 4 show HEVC (HM-14.0 and HM-15.0) vs. H.264/AVC (JM-18.6) video in two different SD resolutions, when Basketball Pass and Party Scene test sequences are processed, respectively.

Fig. 3. HEVC (HM-14.0 and HM-15.0) vs. H.264/AVC (JM-18.6) subjective video assessment for Basketball Pass test sequence.

Fig. 4. HEVC (HM-14.0 and HM-15.0) vs. H.264/AVC (JM-18.6) subjective video assessment for Party Scene test sequence.

Fig. 5 and Fig. 6 show HEVC (HM-14.0 and HM-15.0) vs. H.264/AVC (JM-18.6) video in HD and full HD resolution, when City (HD) and Kimono1 (full HD) test sequences are processed, respectively. In the left part of the window, all tested test sequences are shown after decoding process in JM 18.6 decoder, while in right part of the same window, same test sequences are shown after decoding process in HM-14.0 and HM-15.0 decoder.

Fig. 5. HEVC (HM-14.0 and HM-15.0) vs. H.264/AVC (JM-18.6) subjective video assessment for City test sequence.

Fig. 6. HEVC (HM-14.0 and HM-15.0) vs. H.264/AVC (JM-18.6) subjective video assessment for Kimono1 test sequence.

Objective test results show that there is negligible loss in the term SNR for luma component of picture when comparing HEVC to H.264/AVC video. On the other hand, when four test sequences in different classes and resolutions are compared on subjective way, it is evident that HEVC reached the same results from subjective point of view.

5 Conclusion

The results presented in this paper indicate that the HEVC standard can provide a significant amount of increased coding efficiency compared to previous H.264/AVC standards. The results of objective tests are presented, where SNR, bit-rate and encoding time saving are measured, when the coding efficiency of the capabilities concerning two different HEVC encoders (HM-14.0 and HM-15.0) vs. H.264/AVC (JM-18.6) encoder are compared. Also, results of subjective tests are provided comparing HEVC (HM-14.0 and HM-15.0) vs. H.264/AVC (JM-18.6). The results indicate that a
bit-rate reduction can be achieved more than 50%, while the encoding time are slightly decreased in average over 8% with negligible loss in term signal-to-noise ratio (SNR) for using test video sequences in different classes and resolutions.

Our study in the next step will focus on encoder algorithm solutions in order to improve performance of the HEVC to be more interesting to the global market and industry.

References: