

Review of Proposed High Efficiency Video Coding (HEVC) Standard

Eskinder Anteneh Ayele
Visvesvaraya National Institute of Technology,
Nagpur, India

S.B.Dhok
Visvesvaraya National Institute of Technology,
Nagpur, India

ABSTRACT

The objective of this paper is to provide a review of the recent developments in HEVC standardization, particularly focusing on the key features of hybrid coding tools and gives theoretical analysis, summaries of the technological advancements, and compares its performance with the H.264/MPEG-4 AVC high-profile standard. High Efficiency Video Coding (HEVC) is a proposed and under development new generation of video compression standard, which enables substantially higher compression capability than the existing state-of-the-art video coding standard and exhibit superior coding performance improvements over its predecessors. Some of the major contributors to the higher compression performance of HEVC are the introduction of quadtree structure, improved techniques to support parallel encoding/decoding, more directional intraprediction modes, support for several integer transforms with square as well as nonsquare transforms, merging of prediction blocks for improved motion information encoding, and extensive In-loop processing on reconstructed pictures. When completed, it unlocks future business not possible with today's AVC notably suitable for resolutions up to Ultra High Definition (UHD) video coding in the future.

General Terms

High Efficiency Video Coding (HEVC)

Keywords

H.264/ MPEG-4 AVC, HEVC, Quadtree structure, Intraprediction modes, nonsquare transforms, In-loop processing.

1. INTRODUCTION & BACKDROUND

With the advent of the multimedia and information era, the past two decades have seen the major developments of video coding standards and video compression technologies in significant compression gains and impressive deployment of services and applications. The research on the basic coding tools involved in video compression actually started as early as in the 1950s and 1960s with spatial differential pulse code modulation (DPCM) coding of images. In the 1970s, transform coding techniques and Motion compensated prediction error coding were investigated, and the well-known block-based discrete cosine transform (DCT) was introduced by Ahmed *et al.* (1974) and (Jain and Jain, 1981; Girod, 1987) [1].

Meanwhile, the introduction of video content has continued to become an increasing presence in our lives, with increasing diversification of usage models along with ever-increasing demands for higher quality. Consumers now expect higher resolution for their entertainment video, as standard-definition analog broadcast TV and VCR movies have given way to

HDTV, DVD, and Bluray, and UHD video has emerged on the prospect.

To meet the industry requirement of standardizing existing video techniques and reassured these issues, standards were developed and running to announce new video compression projects by the two premier international standardization organizations, ISO/IEC and ITU-T. To face the new challenges, the limitations of current technologies prompted the two main video coding standardization players to establish the Joint Collaborative Team on Video Coding (JCT-VC) in January 2010, with the objective to develop and launch new high-efficiency video coding (HEVC) standard, which was pretended to be the natural evolution of H.264/AVC. Then, they issued a joint formal Call for Proposals (CfP) in April 2010. In the necessity document, the main goals of HEVC were stated as coding performance on high resolution pictures, picture size up to 8Kx4K, low delay, and low complexity [2].

In February, totally 27 proposals were submitted to JCT-VC and the subjective image quality evaluation was done in March. The evaluation results discussed in the April JCT-VC meeting at Dresden, Germany [2] showed that some proposals could reach the same visual quality as AVC at only half the bit rate in many of the test cases, at the cost of 2x to 10x increase in computational complexity; and some proposals achieved good subjective quality and bit rate results with lower computational complexity than the reference AVC High profile encodings. At present, this new standard is currently under its final standardization stage and the Final Draft International Standard is expected to be delivered in 2013. Its name will be *High Efficiency Video Coding* (HEVC), but it is also known as *H.265* and *MPEG-H Part 2* [3][4][5]. It is expected to satisfy the ever increasing requirements for cost effective video encoding process by optimizing the video quality, the compression efficiency, the spatial and temporal resolution, and finally the computational complexity.

Finally, the goal of this review paper is to present and highlight the latest developments and analysis of HEVC related technologies and coding tools, and to provide and/or lead readers with a deep understanding of this emerging video coding standard and the related state-of-the-art technologies. Upon this introductory section, the rest of the paper is organized as follows: Section 2- the main body of this paper, introduces the main novel features of HEVC and related work on the development of HEVC coding tools and dedicated to the six basic building blocks. Section 3- presents the performance evaluations through CfP, which have been carried out to be able to appreciate the computational cost of the coder and we show some comparisons with respect to the coding efficiency and complexity between HEVC and the H.264/AVC standards. Section 4- gives the potential

applications and future services of HEVC. At the end, Section 5 concludes the paper.

2. FEATURES/HEVC CODING TOOLS

This paper will focus on the block-based MC/DCT hybrid coding of the residual schemes and give a review of the current design of HEVC and discuss the features that differentiate it from its predecessor. Although HEVC has not yet been finalized, the key elements of this new standard have been identified and it is still being slightly fine-tuned, and will include other features by the time it reaches its final form. It is important to note that this paper serves as a snapshot of the current condition of HEVC as it gets close to its completion status. In that respect, the final version will differ somewhat from what is described.

HEVC is actually based on the same hybrid spatial-temporal prediction system as its predecessor H.264/AVC [7]. Figure 1 shows the block diagram of the basic HEVC encoder design—as it is implemented in the HM 5.0 software codec [7]. As can be observed, the main structure of the HEVC encoder looks like that of the H.264/AVC.

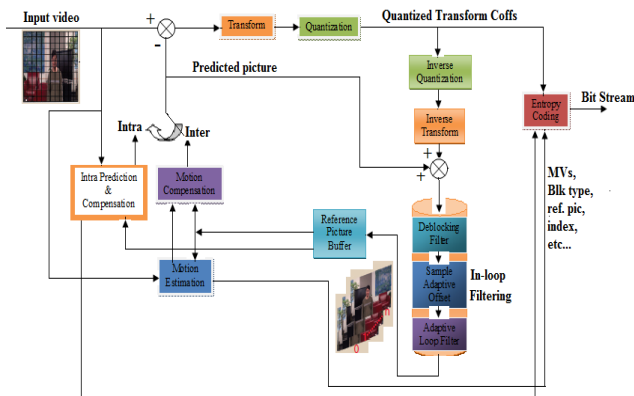


Figure 1: Block diagram of an HM-5.0 encoder

The new tool features are summarized in tabular form in [2] as a comparison with AVC and the main key features of the latest version of HEVC can be reviewed as follows:

- Wide-range variable block-size prediction, with block sizes ranging from 64x64 down to 8x8 pixels and the size of prediction blocks can be adaptively chosen by using recursive quad-tree partitioning [8].
- The block sizes used for DCT based residual coding is adapted to the characteristics of the residual signal by using a nested, so-called residual quadtree (RQT) for partitioning of a given prediction block [9]. It support for several integer transforms, ranging from 32x32 down to 4x4 pixels, as well as nonsquare transforms [7].
- Neighboring blocks can be merged into one region, such that motion information has to be transmitted only once for a whole region [10].
- Interpolation of fractional-sample positions for motion-compensated prediction is based on a fixed-point implementation of the Maximal-Order-Minimum-Support (MOMS) algorithm using an IIR/FIR filter [11].
- Improved mechanisms to support parallel encoding and decoding, including tiles and wavefront parallel processing (WPP) [7].
- Extensive in-loop processing on reconstructed pictures: In addition to the deblocking filter, sample adaptive offset (SAO) and adaptive loop filtering (ALF), a separable Wiener filter is applied within the coding loop [12].

- More intraprediction modes (most of which are directional), which can be done at several block sizes [7].
- The novel probability interval partitioning entropy (PIPE) coding scheme provides the coding efficiency and probability modeling capability of context based adaptive binary arithmetic coding (CABAC) [7].

In addition to the novel algorithms mentioned above, the HEVC proposals also includes the following features [13]; Simple quad-tree structure supporting large macroblock sizes of 32x32 and 64x64, Low complexity B pictures that only use integer motion vectors for SKIP and DIRECT modes, and Spatially Varying Transform (SVT) that allows the position of the transform change arbitrarily within the macroblock.

The main aspects of the proposed coding tools could be described in detail in the subsequent sections.

2.1 Quadtree-Based block Partitioning

An important difference of HEVC compared to H.264/AVC is the frame coding structure. In HEVC each frame is divided into the basic processing unit scheme called Largest Coding Units (LCUs) [8]. LCUs can be recursively split into smaller Coding tree Units (CUs) using a generic quadtree segmentation structure (a nested quadtree structure) that indicates the subdivision of the CU for the purpose of prediction and residual coding. CUs can be further split into Prediction Units (PUs) used for intra- and inter-prediction and Transform Units (TUs) defined for transform and quantization (see figure 2).

However, in H.264/AVC, each picture is partitioned into 16x16 macroblocks, and each macroblock can be further split into smaller blocks (as small as 4x4) for prediction [7]. As the picture resolution of videos increases from standard definition to HD and beyond, the chances are that the picture will contain larger smooth regions, which can be encoded more effectively when large block sizes are used.

This is the reason that HEVC supports larger encoding blocks than H.264/AVC, while it also has a more flexible partitioning structure to allow smaller blocks to be used for more textured and in general uneven regions. Hence, it has been designed to target ultra high resolution with higher frame rates compared to H.264/AVC. Taking this into consideration, HEVC has introduced a new partitioning image scheme concept based on a quadtree structure with larger block size – a 64x64 Coding Unit (CU) and can be recursively further split into 4 CUs (Quadtree), which are used as the basic unit for intra- and inter-coding. The size of CUs can be as large as that of LCUs or and become as small as 8 x 8, depending on the picture content. Because of recursive quarter-size splitting, a content-adaptive coding tree structure comprised of CUs is created in HEVC [14].

Each CU may contain one or more PUs, and each PU can be as large as their root CU or as small as 4x4 in luma block sizes [10]. While an LCU can recursively split into smaller and smaller CUs, the splitting of a CU into PUs is nonrecursive (it can be done only once). PUs can be symmetric or asymmetric. Symmetric PUs can be square or rectangular (nonsquare) and are used in both intraprediction (uses only square PUs) and interprediction. In particular, a CU of size 2Nx2N can be split into two symmetric PUs of size Nx2N or 2NxN or four PUs of size NxN as shown in figure 2b. Asymmetric PUs is used only for interprediction. This allows partitioning, which matches the boundaries of the objects in the picture [14].

Finally, since HEVC applies a DCT-like transformation to the residuals to decorrelated data, TU is the basic unit for transform and quantization, which may exceed the size of PU, but not that of the CU. Only two TU modes are considered [6], signaled by transform unit size flag: i) If the Transform unit size flag = 0 \rightarrow $2N \times 2N$ (i.e., the same as the CU size), ii) Else if the Transform unit size flag = 1 \rightarrow Square units of smaller size are considered: $N \times N$ if PU splitting is symmetric or $N/2 \times N/2$ if PU splitting is asymmetric. So, the size and the shape of the TU depend on the size of the PU. The size of square-shape TUs can be as small as 4×4 or as large as 32×32 . Nonsquare TUs can have sizes of 32×8 , 8×32 , 16×4 , or 4×16 luma samples. Each CU may contain one or more TUs and each square CU may split into smaller TUs in a quad-tree segmentation structure. Figure 2(a-c) shows some examples of partitioning in HEVC.

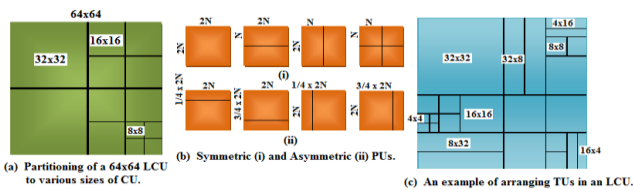


Figure 2: Block partitioning in HEVC

For transform sizes larger than 8×8 , the proposal [15] utilizes truncated transforms where only the 8×8 low frequency coefficients are calculated. This results in a significant computational complexity saving, and implies that only 4×4 and 8×8 quantization kernels are used. The encoder selects the best transform size and position by searching several candidates using rate-distortion optimization [15]. The main idea for supporting variable block-size transforms is to adapt the transform to the varying space-frequency characteristics of the residual signal. Other Features like a novel technique called Spatially Varying Transform (SVT), where the position of the transform block within the macroblock is not fixed but can be varied [15] [16].

HEVC introduced tiles as a means to support parallel processing, with more flexibility than normal slices in H.264/AVC but considerably lower complexity than flexible macroblock ordering (FMO). Tiles are specified by vertical and horizontal boundaries with intersections that partition a picture into rectangular regions [7]. To support parallel processing, each slice in HEVC can be subdivided into smaller slices called entropy slices. Each entropy slice can be independently entropy decoded without reference to other entropy slices. Therefore, each core of a CPU can handle an entropy-decoding process in parallel. However, they both come with a performance penalty since prediction dependencies are broken across boundaries and the statistics used in entropy coding have to be initialized for every slice/tile. To avoid these problems, Wavefront Parallel Processing (WPP) is supported in HEVC [8][14]. Wavefront processing is a way to achieve parallel encoding and decoding without breaking prediction dependencies and using as much context as possible in entropy encoding.

Further, the detail of Quadtree-Based Block Partitioning of HEVC is given in [7] [8].

2.2 Inter and Intra-Prediction Coding

Predictive coding is the primary tool employed in current video compression technologies, and it is very efficient for removing the correlation between pixels in both spatial and temporal domain. Pixel values to be coded are predicted from

already coded and reconstructed adjacent pixel values, and only small prediction errors are coded and the bit rate reduction depends on the correlation coefficient. In HEVC, for generating the prediction signal at subsample positions, one of the proposals [8] is used an interpolation method based on families of so-called maximal-order-minimal-support (MOMS) basis functions. This requires, however, an additional pre-filtering step on the reference picture before the actual interpolation. More details on the subsample interpolation scheme can be found in [8] [11].

2.2.1 Inter-Prediction Coding

Inter-prediction explores temporal redundancy between frames to save coding bits. By using motion compensated prediction, the best matching position of current block is found within the reference picture so that only prediction difference needs to be coded.

The enhancements of inter-prediction introduced in HEVC, compared with H.264/AVC, are described as variable PU size motion compensation, improved subpixel interpolation, and motion parameter encoding and improved skip mode [7].

Each PU coded using interprediction, has a set of motion parameters, which consists of a motion vector, a reference picture index, and a reference list flag. Inter-coded CUs can use symmetric and asymmetric motion partitions (AMPs). AMPs allow for asymmetrical splitting of a CU into smaller PUs. AMP can be used on CUs of size 64×64 down to 16×16 and improves the coding efficiency since it allows PUs to more accurately conform to the shape of objects in the picture without requiring further splitting [17].

The existing sub-pel interpolation method has been improved by replacing the fixed filters by the adaptive ones or by redesigning the filter coefficients [2]. Several proposals adaptively update interpolation filters by the least squares method in order to minimize the prediction errors of each video frame. In [2][18][19][20], multiple sets of filters are transmitted for an adaptive selection at slice or partition level. The extra overheads are reduced by making use of the symmetry properties of these filters. In addition to adjusting filters on the fly, some redesigned filters are proposed. The schemes in [18][21] increase the precision for filtering operations.

The accuracy of motion compensation in HEVC is $1/4$ pel for luma samples. To obtain the noninteger luma samples, separable one-dimensional eight-tap and seven-tap interpolation filters are applied horizontally and vertically to generate luma half-pel and quarter-pel samples, respectively [22]. Moreover, [23] suggests extending the macroblock size up to 64×64 so that new partition sizes 64×64 , 64×32 , 32×64 , 32×32 , 32×16 , and 16×32 can be used. Instead of using the fixed interpolation filter from H.264/AVC, Adaptive Interpolation Filters (AIF) is proposed [23], such as 2D AIF, Separable AIF, Directional AIF, Enhanced AIF, and Enhanced Directional AIF.

To further improve inter prediction efficiency, finer fractional motion prediction and better motion vector prediction were proposed. Increasing the resolution of the displacement vector from $1/4$ -pel to $1/8$ -pel to obtain higher efficiency of the motion compensated prediction is also suggested in [8]. In HEVC, MVs can be predicted either spatially or temporally [24]. Furthermore, HEVC introduces a technique called motion merge. For every inter-coded PU, the encoder can

choose between using explicit encoding of motion parameters, motion merge mode, or the improved SKIP mode [8]. In [2], introduces a parametric overlapped block motion compensation (POBMC) technique to improve inter-frame prediction. For more details on the block merging concept, the reader is referred to [10].

2.2.2 Intra-Prediction Coding

The H.264/AVC intra prediction tool provides DC and several directional modes for predicting variable-size blocks. The predictor is linearly generated from target block's neighboring L-shaped coded pixels. However, this prediction scheme has several inherent weaknesses as: Poor performance inevitably invites when the distances between the reference and the predicted pixels increase, the straightforward design of extrapolation filters is incapable of synthesizing periodical and complex textures, and artificial edges appear along the directions of intra prediction.

Based on the above investigations, many tools like Line-based Prediction [25][26], Pyramid and Interleaved Prediction [25] [27], Template Matching Average [2][19], Pre- and Post-filtering [24][28], Plane Prediction [21], Chroma prediction [28], and Extended Directional Prediction [21][28] are proposed to alleviate these problems. The bi-directional intra prediction (BIP) [30] also another proposal, deduces the predictor from averaging the prediction results of two different modes.

HEVC uses block-based intraprediction to take advantage of spatial correlation within a picture and it follows the basic idea of H.264/AVC intraprediction but makes it far more flexible. HEVC has 35 luma intraprediction modes compared with nine in H.264/AVC. Furthermore, intraprediction can be done at different block sizes, ranging from 4x4 to 64x64. Figure 3 shows the luma intraprediction modes of HEVC and H.264. HEVC also includes a planar intraprediction mode, which is useful for predicting smooth picture regions. In planar mode, the prediction is generated from the average of two linear interpolations (horizontal and vertical) [21].

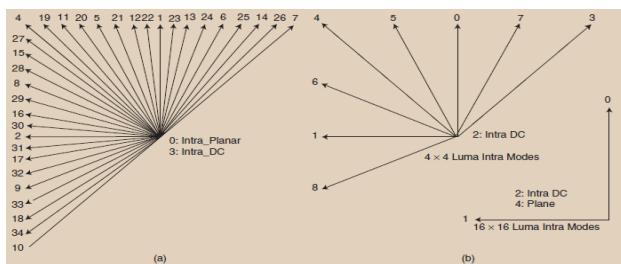


Figure 3: Luma intraprediction modes of (a) HEVC and (b) H.264/AVC

In general, the current intra prediction technique in HEVC unifies two simplified directional intra prediction methods [6]: the Arbitrary Direction Intra and the Angular Intra Prediction. To improve the performance of intra-prediction, mode dependent intra smoothing (MDIS) is used for some intra-modes which involves applying a simple low-pass FIR filter with coefficients (1, 2, 1)/4 to the samples being used for prediction. See Figure 9 of [7] for detail and an example of an MDIS application.

The proposal [15] introduces two techniques (angular prediction and planar coding) in more detail to improve the

visual quality of decoded video, both of which are included in the HEVC TMuC.

2.3 Transform Coding

In H.264/AVC and MPEG-4 standards, the DCT basis is not optimal for various directional patterns in residual signals [2]. The transform basis should be made adaptable to the statistical variation of realizations. Therefore, anticipation of a need for better transform coding tools leads to redesigning the existing DCT-based coding for further optimizing the energy compaction of residual signals.

HEVC applies a DCT-like integer transform on the prediction residual. HEVC includes transforms that can be applied to blocks of sizes ranging from 4x4 to 32x32 pixels and also supports transforms on rectangular (non-square) blocks particularly in case of Asymmetric Motion Partitioning (AMP), Non-Square Transform (NSQT) is used [13]. The integer transforms used in HEVC are better approximations of the DCT than the transforms used in H.264/AVC. The basis vectors of the HEVC transforms have equal energy and there is no need to compensate for the different norms, as in H.264/AVC. In HEVC, the splitting of CU into smaller square TUs is signaled using a second quadtree, the residual quadtree (RQT) [9], which allows adapting the transform to the frequency characteristics of the residual signal.

In many proposals [2] [13], mode dependent directional transform (MDDT) and rotational transform (ROT) are widely used since they have been proven to be effective for decorrelating the redundancies along the directions of intra prediction. In MDDT [2], each intra prediction mode is coupled with a unique pair of transform matrices, which is derived from the off-line training processes of Karhunen Loéve transform (KLT), for the strongly mode-dependent residual signals. ROT chooses to change the DCT basis rather than to train a new KLT basis. The energy of residual signals is generally concentrated on low-frequency bands after the DCT.

HEVC also incorporates a 4x4 discrete sine transform (DST), which is used for blocks coded with some directional intra-prediction modes. When using intra-prediction, the pixels close to the ones used for prediction (i.e., near the top or left boundaries) will usually be predicted more accurately than the pixels further away. Therefore, the residuals tend to be larger for pixels away from the boundaries. The DST will usually be better at encoding these kinds of residuals, because the DST basis functions start low and increase, compared with the DCT basis functions that start high and decrease [7].

2.4 Quantization

In the MC/DCT hybrid video coding schemes, uniform scalar quantization schemes are usually utilized to quantize the transform coefficients, and the quantization step size, which determines the quantization strength, is transmitted to the receiver. To achieve better quantization, optimized quantization decision at the macroblock level and at different coefficient positions are proposed. More recently, for HEVC [23] gives an improved, more efficient Rate Distortion Optimized Quantization (RDOQ) implementation. In [23], Adaptive Quantization Matrix Selection (AQMS), a method deciding the best quantization matrix index, where different coefficient positions can have different quantization steps, is proposed to optimize the quantization matrix at a macroblock level. The quantization weighting matrix, which is controlling

element can be either uniquely defined and sent to the decoder as coding parameters, or substituted by a default one. To match the statistics of the transform coefficient distribution, adaptive selection of the quantization weighting matrix is proposed in [2].

2.5 Coding

After transformation and quantization, entropy coding is applied to code all the syntax elements and quantized transform coefficients in video coding techniques. Even though context-adaptive binary arithmetic coding (CABAC) is proven to be efficient in H.264/AVC [2], it is designed for serial processing and its context adaptive feature is based on the statistics of previously coded data. A low data throughput is unavoidable and becomes a bottleneck on handling high resolution videos. Therefore, a new design for entropy encoder should consider parallelism, load-balance and complexity/performance transactions.

So-far, the parallel processing capabilities of CABAC are improved in entropy-slice-level, syntax-level and bin-level parallelism aspects [24] [28]. [21] Proposes a context-adaptive variable length coding (CAVLC) design for both residual and non-residual information with two major features. One is to improve the coding efficiency by providing more VLC tables. The other is to improve the context adaptively by maintaining a sorting table.

The HEVC defines two context-adaptive entropy coding patterns [6], one for the higher complexity mode and one for the lower-complexity mode. The lower-complexity mode is based on a variable length code (VLC) table selection for all the syntax elements, while the selection of a particular code table is done in a context-based scheme, depending mainly on the previous decoded values. To improve the worst-case throughput, the codec uses a higher-throughput alternative mode for coding transform coefficient data. Figure 12 in [7] illustrates the block diagram of HEVC entropy coding. Using this harmonized design, HEVC entropy coding uses the best features of both CABAC and CAVLC coding (i.e., high efficiency and low complexity, respectively).

Some techniques currently under consideration for the HEVC standard also include: Probability Interval Partitioning (PIPE) and Syntax Element Partitioning (SEP) [30] [31]. PIPE uses an entropy coding algorithm similar to CABAC in H.264/AVC. The main difference is in the binary arithmetic coder in which instead of coding the bins using a single arithmetic coding engine a set of encoders are used each one associated to a partition of the probability interval. SEP consists of grouping bins in a slice by the type of syntax element rather than by macroblock (or LCU) as is in H.264/AVC.

For reducing the rounding errors in intra prediction (transform and in-loop filtering processes), the internal bit-depth increasing (IBDI) proposals [29] [19] increases the calculation precision during the coding process. Furthermore, [23] proposed to use Decoder Side Motion Estimation (DSME) for B-picture motion vector decision, which improves coding efficiency by saving bits on B-picture motion vector coding.

2.6 In-loop Processing/Filtering

In HEVC internal scheme, two types of filtering processes have been obtained. The first one is Interpolation Filtering (IF), which is used to obtain the samples at fractional pixels

for motion vectors pixel accuracy and the second one is In-Loop Filtering (LF), in order to restore the degraded frame caused by compression.

2.6.1 Interpolation Filtering (IF)

Similar to H.264/AVC, the proposals for HEVC utilizes a translational motion model with motion vectors having quarter pixel accuracy. Here to obtain the samples at fractional pixels, two sets of interpolation filters are utilized. The first set of interpolation filters is referred to as the Directional Interpolation Filter (DIF) and is used for all the quarter-pixel positions. Directional filters use either 6-tap or 12-tap filter for each sub-pixel, and has significantly less complexity compared to H.264/AVC interpolation, where 6x6 separable filters are utilized [23]. The second set of interpolation filters consists of Separable Filters (SF) where interpolated samples are calculated by first applying a 6-tap filter horizontally and then vertically. More detail related to interpolation filtering is given in [15].

2.6.2 In-Loop Filtering (LF)

The name loop filtering reflects the fact that filtering is done as part of the prediction loop rather than post-processing. HEVC expands an in-loop processing by introducing two new tools: Sample Adaptive Offset (SAO) and Adaptive Loop Filter (ALF), mainly aims at reducing additional quantization noise in the output of the deblocking filter. Both types of filter are highly adaptive, and they are both applied within the coding loop with the output of the final filtering stage being stored in the reference picture buffer.

2.6.2.1 Deblocking Filter

To improve the subjective and objective quality of the video by reducing blocking artifacts, H.264/AVC uses low-pass filters which are adaptively applied to each decoded picture, before it goes into the decoded picture buffer at block boundaries. The filter strength is adaptively adjusted according to the boundary strength.

In HEVC, there are several kinds of block boundaries, such as CUs, PUs, and TUs). Filtering is potentially performed on all block edges down to the 8x8 block level. Transform block edges on the 4x4 level are not filtered. This reduces filter complexity without compromising the subjective quality for higher resolutions [7] [15]. For each boundary, a decision is made to turn the deblocking on/off and whether to apply strong or weak filtering. Since only half the number of block edges are filtered, and due to less complex logic for enabling/disabling the filter, it has a lower computational complexity than the H.264/AVC deblocking filter. The proposals in [18] [26] [32] simplify the deblocking filter complexity.

Another beneficial feature as a new approach called “parallel deblocking filter” is introduced in [33]. In this scheme the deblocking filter is divided in two separated frame stages: horizontal and vertical filtering.

2.6.2.2 Sample Adaptive Offset filter (SAO)

The sample adaptive offset (SAO) filter is applied in between the deblocking filter and the ALF. It is a new coding tool introduced in HEVC, which involves classifying pixels into different categories based on either intensity or edge properties and adding a simple offset value either band offset (BO) or edge offset (EO), to the pixels in each category in a region to reduce distortion.

Band offset (BO) classifies all pixels of a region into multiple bands, with each band containing pixels in the same intensity interval [15]. The intensity range is divided into 32 equal intervals from zero to the maximum intensity. The 32 bands are divided into two groups and only offsets in one group are transmitted. The encoder decides which group of bands to apply SAO, so 16 offsets will be encoded in the bit stream [34].

Edge offset (EO) uses four 1-D 3-pixel patterns for pixel classification with consideration of edge directional information, as illustrated in Figure 4. Each region of a picture can select one pattern to classify pixels into multiple categories by comparing each pixel with its two neighboring pixels. Each pixel can be classified as a peak (if it is greater than two neighbors), valley (if it is less than the two neighbors), edge (if it is equal to one neighbor, categories 2 and 3), or none of these. Four offset values will be calculated for these four categories [7]. The encoder can choose to apply either BO or EO to different regions of a picture and also signal that neither BO nor EO is used for a region.



Figure 4: EO pixel classification patterns

2.6.2.3 Adaptive Loop Filter (ALF)

ALF is a quadtree-based separable 2D Wiener filter which allows minimizing the MSE between the reconstructed and source frames. It classifies the pixels into multiple categories based on the local directional characteristics. A Wiener filter is estimated and transmitted for each category by minimizing the average mean square error between the original and reconstructed pixels in that category [13]. The filter is adaptive in the sense that the coefficients are signaled in the bit stream and can therefore be designed based on image content and distortion of the reconstructed picture.

Two different ALF techniques are adopted so far: Quadtree-based Adaptive Loop Filter (QALF) and Block-based Adaptive Loop Filter (BALF) [23]. The main idea of Quadtree ALF (QALF) is to signal the on/off decision of filtering through a quad-tree partition process. QALF is adopted, and improved by providing multiple filters for adaptation, as suggested by many HEVC proposals [18][19][20][26].

The ALF is LCU independent and can be executed for each block in parallel. In proposal [30], to reduce cache line conflicts and synchronization overhead, eight consecutive LCUs are grouped in a work unit and processed by a single core. The current HEVC draft [14] uses a single filter shape, a cross overlaid on a 3x3 square with nine coefficients to be encoded in the bit stream shown in [35]. Note that the number of taps in the filter is greater than nine due to symmetry.

There are two modes that can be used for applying different filters to different pixels within each picture: region based adaptation (RA) and block-based adaptation (BA) [7]. The filter coefficients for each region can be calculated based on the autocorrelation and cross-correlation of the original pixels and the reconstruction pixels in the region (using Wiener-Hopf equations) [36]. The ALF can be enabled or disabled for different picture areas based on the partitioning of LCUs into

CUs (in a quad-tree segmentation structure). More details is given in [2][6][7][8][12][36].

3. PERFORMANCE ANALYSIS AND COMPARISONS

In this section the review of the compression and encoding efficiencies of the HEVC algorithms in comparison to the H.264/AVC for the test signals under test used in some papers, when same encoding parameters have been selected. The section also considers and presents the experimental results of different proposals for the comparison.

From the previous sections, the summary of the main factors of HEVC for coding gain over H.264/AVC and factors that leading to reduced complexity in tabular form as follows:

Table 1. Summary of the main factors of HEVC for coding gain and reduced complexity over H.264/AVC

Main factors for coding gain over H.264/AVC	Main factors leading to reduced complexity
<ul style="list-style-type: none"> - More flexible temporal prediction and scanning structure. - Larger and more flexible coding, prediction, and transform units. - Improved mechanisms to support parallel encoding & decoding. - More accurate intra prediction approach and directions/modes. - More accurate motion parameters (including merge mode) and sub-pixel prediction. - Inclusion of non-square transform and allowing asymmetric motion prediction. - More flexible transform, choice of DST, and no-transform option. - Rate-distortion optimized quantization (RDOQ). - Improved in-loop filters, including the new sample adaptive offset filter (SAO). - Use of mean opinion score (MOS) as visual quality metric. 	<ul style="list-style-type: none"> - Average interpolation complexity is less than H.264/AVC due to DIF. - Memory bandwidth consumption due to motion compensation is lower than H.264/AVC, because motion partitions smaller than 8x8 are not used. - The interpolation complexity for B pictures is significantly less than in H.264/AVC as the SKIP and DIRECT modes use motion vectors rounded to integers. - The deblocking filtering has significantly lower computational complexity due to smaller number of edges being checked. - The VLC (de)coding of coefficients is simpler than H.264/AVC CALVC.

The technical assessment of the HEVC technology, as it was performed in several JCT-VC meetings, revealed that all proposed algorithms were based on the traditional hybrid coding approach, combining motion compensated prediction between video frames with intra-picture prediction, closed-loop operation with in-loop filtering, 2D transformation of the spatial residual signals, and advanced adaptive entropy coding. Many specific candidate technology improvements were identified from the proposal responses, as was summarized in technology survey documents [37].

As an initial step toward moving forward into collaborative work, the *Test Model under Consideration* (TMuC) document [38] was produced, combining identified key elements from a group of seven well-performing proposals. This first TMuC became the basis of a first software implementation, which after its development has begun to enable more rigorous assessment of the coding tools that it contains as well as additional tools to be investigated within a

process of “Tool Experiments” as planned at the first JCT-VC meeting.

In order to evaluate the performance of HEVC, and to compare it with the performance of H.264/AVC, a set of simulations have been launched, whose conditions have been chosen according to JCT-VC documents. As basic test conditions used for the CfP and refined in subsequent work, they considered five classes of source video test sequences as specified in [5] and tabulated in [3]. Similarly, three basic use cases have been established as representative of relevant applications for HEVC CfP testing as: Random Access (RA), in which intra refresh is relatively frequent and delay is not especially critical, Low Delay (LD), for real-time communication with minimal delay and without random-access intra refresh, and All-Intra (AI), in which no inter-picture temporal prediction is applied.

For testing HEVC, the random access high efficiency (RA-HE) configuration was used [5], to ensure achieving the highest compression performance. The RA-HE configuration is as follows: hierarchical B pictures, GOP length 8, ALF, SAO, and RDOQ were enabled (see [14] and [39] for more details). Figure 16 in [7] shows the rate-distortion (RD) curves for all the test sequences, and Table 4 in [7] lists the average PSNR improvement and average PSNR savings achieved by HEVC over the H.264/AVC standard. As it can be observed, the current HEVC design outperforms H.264/AVC by 29.14–45.54% in terms of bit rate or 1.4–1.87 dB in terms of PSNR. Subjective comparison of the quality of compressed videos—for the same (linearly interpolated) mean opinion score points—shows that HEVC outperforms H.264/AVC, yielding average bit-rate savings of 58%.

Similarly, [39] presents the results of objective (PSNR-based) compression comparison tests between the current state of the HEVC draft standard and the AVC High Profile as an anchor reference. The performance is measured with all HEVC tools activated except 10-bit encoding for the all-intra, the random-access, and the Main Profile low-delay configurations. As a result the HM-6.0 (HEVC reference software) can save about 24%, 36%, and 37% in bit rates respectively.

In the CfP [40], two sets of application-specific coding conditions were specified: so-called constraint set 1 (CS 1) as a random access configuration (e.g., for broadcast applications) and constraint set 2 (CS 2) as a low-delay configuration (e.g., reflecting real-time applications).

In [8] for both constraint sets, they configured the encoder to use a maximum prediction block edge length of $N_{\max} = 64$ luma samples and a maximum prediction quadtree depth of 4, corresponding to $N_{\min} = 4$. Table I in [11] shows the average bit rate savings of the bit streams as submitted to the CfP with the average for each of the test sequences. Overall, significant objective gains in terms of average 29.6% BD rate savings relative to the H.264/AVC HP anchor have been achieved for the random access scenario, and 22.68% BD rate savings for the low delay scenario.

In order to test the encoding and decoding complexity of the proposed algorithm, extensive simulations have been conducted. The simulation results show that the encoding of the proposed algorithm is around 25 times faster than JM17.0 encoding, and decoding is two to three times faster than that of JM17.0 [15]. As it can be observed from the results of Table 2 of [6], the HEVC encoded signals appear to have

approx. 50% better encoding efficiency in comparison to H.264/AVC. The results of this table verify the objective of the HEVC that aims to double the compression ratio in comparison to the AVC has been achieved.

4. APPLICATIONS

Regarding the compression gain already demonstrated, one cannot deny that HEVC will offer technical and commercial benefits to existing applications and usage scenarios. Hence, potential applications lie on IPTV (SD or HD) over DSL, where HEVC would increase service reach, or on point-to-point contribution on premium or light links. Moreover, multiscreen applications or over-the-top (OTT) services can also benefit from HEVC by improving the overall quality of video to mobile devices. On the other hand, HEVC enables future services not possible with today’s state-of-the-art compression standards. Hence, potential applicable uses would be delivering 1080p60/50 at bitrates comparable to today’s 1080i data rates, full resolution HD 3DTV at today’s HD delivery rates. HEVC enabling better picture quality at lower bitrates, it will offer sports fans, for instance, a better viewing experience on mobile devices.

Last but not least, a 4K application for which HEVC was originally designed. Indeed, the broadcast industry is completing the transition to HD all around the world, and demand for an even enhanced end-user experience especially for home theater and premium events like sports is rising. 4KTV would be the answer and HEVC the straightforward video compression technology involved for 4K delivery to the home.

5. CONCLUSIONS

In this paper different topical papers have surveyed and presented as a review of the recent proposals for the High Efficiency Video Coding (HEVC) standardization project. It follows the same basic design principles as already established by the state-of-the-art video coding standard H.264/AVC high profile, with generalizations for more flexibility in terms of prediction and transform block partitioning. The new features of the proposals for HEVC design included the merging of prediction blocks, the MOMS based interpolation scheme, 1-D directional interpolation filters, angular prediction, planar coding, truncated transform techniques, low complexity deblocking filter, the adaptive in-loop filtering, and the Probability Interval Partitioning Entropy (PIPE) coding. We have seen that many compression improvements were made by increasing adaptability.

Most of objective and subjective results from the reviewed CfPs confirm that the goal of developing a high-efficiency video coding standard, which delivers significantly improved compression performance relative to that of the H.264/AVC, and will soon lead to the definition of a new video compression standard developed by JCT-VC with only half of the bit rate.

6. REFERENCES

- [1] Lu YU, Jian-peng WANG, “Hot topic: Review of the current and future technologies for video compression”, Journal of Zhejiang University-SCIENCE C (Computers & Electronics), 2010 11(1):1-13
- [2] Hsueh-Ming Hang, Wen-Hsiao Peng, Chia-Hsin Chan and Chun-Chi Chen, “Towards the Next Video Standard: High Efficiency Video Coding”, Proceedings of the

- Second APSIPA Annual Summit and Conference, pages 609–618, Biopolis, Singapore, 14-17 December 2010.
- [3] Bin Li, Gary J. Sullivan, and Jizheng Xu, "Compression Performance of High Efficiency Video Coding (HEVC) Working Draft 4", pp-886-889 IEEE, 2012.
- [4] JCT-VC, "Report of Subjective Test Results of Responses to the Joint Call for Proposals (CfP) on Video Coding Technology for High Efficiency Video Coding (HEVC)," Document JCTVC-A204, Dresden, DE, Apr. 2010.
- [5] F. Bossen, "Common test conditions and software reference configurations," Document JCTVC-F900, Torino, IT, July 2011.
- [6] Harilaos KOUMARAS, Michail-Alexandros KOURTIS, Drakoulis MARTAKOS, "Benchmarking the Encoding Efficiency of H.265/HEVC and H.264/AVC", IIMC International Information Management Corporation, 2012, ISBN: 978-1-905824-30-4
- [7] Mahsa T. Pourazad, Colin Doutre, Maryam Azimi, and Panos Nasiopoulos, "HEVC: The New Gold Standard for Video Compression", IEEE CONSUMER ELECTRONICS MAGAZINE, JULY 2012.
- [8] Detlev Marpe, Heiko Schwarz, Thomas Wiegand, Sebastian Boße, Benjamin Bross, Philipp Helle, Tobias Hinz, Heiner Kirchhoffer, Haricharan Lakshman, Tung Nguyen, Simon Oudin, Mischa Siekmann, Karsten Sühring, and Martin Winken, "Improved Video Compression Technology and the Emerging High Efficiency Video Coding Standard", IEEE International Conference on Consumer Electronics - Berlin (ICCE-Berlin), 2011.
- [9] M. Winken, P. Helle, D. Marpe, H. Schwarz, and T. Wiegand, "Transform coding in the HEVC test model," in *Proc. IEEE International Conference on Image Processing (ICIP)*, Sep. 2011.
- [10] S. Oudin, P. Helle, J. Stegemann, C. Bartnik, B. Bross, D. Marpe, H. Schwarz, and T. Wiegand, "Block merging for quadtree-based video coding," in *Proc. IEEE International Conference on Multimedia and Expo (ICME)*, Jul. 2011.
- [11] H. Lakshman, B. Bross, H. Schwarz, and T. Wiegand, "Fractional-sample motion compensation using generalized interpolation," in *Proc. Picture Coding Symposium (PCS)*, Dec. 2010.
- [12] M. Siekmann, S. Boße, H. Schwarz, and T. Wiegand, "Separable Wiener filter based adaptive in-loop filter for video coding," in *Proc. Picture Coding Symposium (PCS)*, Dec. 2010.
- [13] Philippe Bordes, Gordon Clare, Félix Henry, Mickaël Raulet, Jérôme Viéron, "An overview of the emerging HEVC standard", IEEE, 2010.
- [14] B. Bross, W. J. Han, J. R. Ohm, G. J. Sullivan and T. Weingand, "High efficiency video coding (HEVC) text specification draft 6," JCT-VC Document, JCTVC-H1003-v21, April 2012.
- [15] Kemal Ugur, Kenneth Andersson, Arild Fuldseth, Gisle Bjøntegaard, Lars Petter Endresen, Jani Lainema, Antti Hallapuro, Justin Ridge, Dmytro Rusanovskyy, Cixun Zhang, Andrey Norkin, Clinton Priddle, Thomas Rusert, Jonatan Samuelsson, Rickard Sjöberg, Zhuangfei Wu, "LOW COMPLEXITY VIDEO CODING AND THE EMERGING HEVC STANDARD", 28th Picture Coding Symposium, PCS2010, Nagoya, Japan, December 8-10, 2010.
- [16] C. Zhang; K. Ugur, J. Lainema, M. Gabbouj, "Video coding using variable block-size spatially varying transforms," IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 905-908, May 2009.
- [17] I.-K Kim, W.-J Han, J. H. Park, and X. Zheng, "CE2: Test results of asymmetric motion partition (AMP)," JCTVC-F379, July 2011.
- [18] M. Karczewicz and et al., "Video coding technology proposal by Qualcomm Inc.," *ISO/IEC JTC1/SC29/WG11, JCTVC-A121*, April 2010.
- [19] I. Amonou and et al., "Description of video coding technology proposal by France Telecom, NTT, NTT DOCOMO, Panasonic and Technicolor," *ISO/IEC JTC1/SC29/WG11, JCTVC-A114*, April 2010.
- [20] Y.-W. Huang and et al., "A Technical Description of MediaTek's Proposal to the JCT-VC CfP," *ISO/IEC JTC1/SC29/WG11, JCTVCA109*, April 2010.
- [21] K. Ugur and et al., "Description of video coding technology proposal by Tandberg, Nokia, Ericsson," *ISO/IEC JTC1/SC29/WG11, JCTVC-A119*, April 2010.
- [22] E. Alshina, A. Alshin, J.-H. Park, J. Lou, and K. Minoo, "CE3: 7 taps interpolation filters for quarter pel position MC from Samsung and motorola mobility," JCTVC-G778, Geneva, Nov. 2011.
- [23] Li Liu, Robert Cohen, Huifang Sun, Anthony Vetro, Xinhua Zhuang, "New Techniques for Next Generation Video Coding", TR2010-058, April 2010.
- [24] M. Winken and et al., "Description of video coding technology proposal by Fraunhofer HHI," *ISO/IEC JTC1/SC29/WG11, JCTVC-A116*, April 2010.
- [25] H. Yang and et al., "Description of video coding technology proposal by Huawei Technologies and Hisilicon Technologies," *ISO/IEC JTC1/SC29/WG11, JCTVC-A111*, April 2010.
- [26] H. Y. Kim and et al., "Description of video coding technology proposal by ETRI," *ISO/IEC JTC1/SC29/WG11, JCTVC-A127*, April 2010.
- [27] K. Sugimoto and et al., "Description of video coding technology proposal by Mitsubishi Electric," *ISO/IEC JTC1/SC29/WG11, JCTVCA107*, April 2010.
- [28] K. McCann and et al., "Samsung's Response to the Call for Proposals on Video Compression Technology," *ISO/IEC JTC1/SC29/WG11, JCTVCA124*, April 2010.
- [29] T. Chujoh and et al., "Description of video coding technology proposal by TOSHIBA," *ISO/IEC JTC1/SC29/WG11, JCTVC A117*, April 2010.
- [30] Mauricio Alvarez Mesa, Chi Ching Chi, Thomas Schierl and Ben Juurlink, "Evaluation of Parallelization Strategies for the Emerging HEVC Standard", 2012.
- [31] Detlev Marpe, Heiko Schwarz, and Thomas Wiegand. Entropy Coding in Video Compression using Probability

- Interval Partitioning. In Picture Coding Symposium (PCS 2010), pages 66{69, Dec. 2010.
- [32] J CT-VC, “Encoder-side description of test model under consideration,” in *Proc. JCT-VC Meeting*, Geneva, Switzerland, July 2010, JCTVC- B204.
- [33] Masaru Ikeda, Junichi Tanaka, and Teruhiko Suzuki. Parallel deblocking filter. Technical Report JCTVC-E181, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T and ISO/IEC, March 2011.
- [34] C.-M. Fu, C.-Y. Chen, Y.-W. Huang, and S. Lei, “Sample adaptive offset for HEVC,” in *Proc. Int. Workshop on Multimedia Signal Processing (MMSP)*, Oct. 2011.
- [35] P. Lai, F. C. A. Fernandes, H. Guermazi, F. Kossentini, and M. Horowitz, “CE8 Subtest 4: ALF using vertical-size 5 filters with up to 9 coefficients,” JCTVC-F303, July 2011.
- [36] C.-Y. Tsai, C.-Y. Chen, C.-M. Fu, Y.-W. Huang, and S. Lei, “One-pass encoding algorithm for adaptive loop filter in high-efficiency video coding,” in *Proc. Visual Communications and Image Processing (VCIP)*, Nov. 2011.
- [37] Gary Sullivan “Meeting report of the ninth meeting of the Joint Collaborative Team on Video Coding (JCT-VC)”, Geneva, CH, 27 Apr – 7 May 2012
- [38] Gary J. Sullivan and Jens-Rainer Ohm, "Recent developments in standardization of high efficiency video coding (HEVC)", SPIE Applications of Digital Image Processing XXXIII, Andrew G. Tescher (editor), Proceedings of SPIE Volume 7798, Paper number 7798-30, August, 2010.
- [39] B. Li, G. J. Sullivan, J. Xu, “Comparison of Compression Performance of HEVC Draft 6 with AVC High Profile,” JCTVC-I0409, JCT-VC, Geneva, 2012.
- [40] N Sprljan “Joint Call for Proposals on Video Compression Technology”, ITU-T Q6/16 document VCEG-AM91 and ISO/IEC JTC1/SC29/WG11 document, N11113, Jan. 2010.