MPEG Video deployment in interactive multimedia systems: HEVC vs. AVC codec performance study

DRAGORAD MILOVANOVIC, ZORAN BOJKOVIC
University of Belgrade
Studentski Trg 1, 11000 Belgrade
SERBIA
dragoam@gmail.com   http://www.zoranbojkovic.com

Abstract: - The coding efficiency of three generations of video coding standards is compared by means of PSNR and subjective testing in interactive video applications, such as video chat, video conferencing and telepresence systems. An unified approach is applied to the analysis of designs of MPEG-H HEVC/H.265, MPEG-4 AVC/H.264, and H.263 reference software implementation. The results of performance tests for selected HD720 video sequences indicate that HEVC encoders achieve equivalent objective video quality as encoders that conform to AVC when using approximately 60% less bit rate on average. Bitrate reduction BRr and coding gain CG based on PSNR measure and complexity based on encoding/decoding time of HEVC MP vs. AVC HP vs. H.263 CHC are tested at bitrates of 0.256, 0.384, 0.512, 0.850, and 1.500 Mbps using the low-delay encoding constraints typical for real-time conversational applications. Low-delay coding are considered by selecting appropriate prediction structures and coding options in software configuration of encoders. Real-time decoding complexity where studied on personal *Ultrabook* x86 computer in conversational application.

Key-Words: - operational configuration, coding gain, subjective quality, codec complexity

1 Introduction

Video coding standards have evolved through the development of the *International Standardization Organization / International Electrotechnical Commission* (ISO/IEC) and *International Telecommunication Union – Telecommunication sector* (ITU-T) standards [1]. The next generation video coding standard HEVC (*High Efficiency Video Coding*) is the joint development project of the ISO/IEC *Moving picture experts group* (MPEG) standardization organization and the ITU-T *Video coding experts group* (VCEG), working together in *Joint collaborative team of video coding* (JCT-VC) [2, 3]. In ISO/IEC, the HEVC standard has become IS 23008-2 (MPEG-H Part 2) while in ITU-T, it is *Recommendation* H.265. HEVC aims to improve coding efficiency (encoded bitrate vs. video quality) in tradeoff computational complexity. The extensions of HEVC support several additional applications scenarios including scalable video coding, and 3D stereo/multiview video coding [4].

The second generation video coding standard directly preceding the HEVC project was AVC/H.264 which was developed in the period 1999-2009 [5, 6, 7, 8]. MPEG-4 AVC (*Advanced Video Coding*) has been an enabling technology for digital video in almost every area that was not previously covered by the first generation MPEG-2/H.262 and H.263 video codecs. AVC is widely used for many applications, including broadcast of DTV formats, video content acquisition and editing systems, camcorders, Internet streaming and mobile network video as well as real-time conversational applications [9].

The primary goal of HEVC/H.265, AVC/H.264 and H.263 digital video coding standards is to optimize coding efficiency. *Coding efficiency* is the ability to minimize the bitrate BR necessary for encoded video content to reach a given level of video quality – or, as alternatively formulated, to maximize the video quality achievable within a given available bitrate. Since all video coding standards of ISO/IEC and ITU-T specify only the bitstream syntax of hybrid transform-entropy encoder with motion-compensated prediction (Fig. 1), and the basic decoding process, they do not guarantee any particular coding efficiency. In our study, all encoders are configured to use the same encoding techniques. The maximum coding delay constraint is considered by selecting appropriate prediction structures and coding options in software configuration of encoders [10].

In the 2nd section of the paper, the main coding tools that contribute to the coding efficiency improvement from one standard generation to the next, are described, followed with low-delay encoder configuration. The coding gain results are presented in the 3rd section, followed with codec complexity test.

2 HEVC vs. AVC vs. H.263 codec

The 1st generation video coding standard H.263 for low bitrate communication was developed in the period 1993-1998 [11]. The first version of ITU-T Rec. H.263 defines syntax features that are very similar to those of H.262 MPEG-2 Video [12], but it includes some changes that make it more efficient for low-delay low bit rate coding. The second and third versions of H.263, which are often called H.263+ and H.263++, respectively, add several optional coding features in the form of annexes. The H.263 profiles that provide the best coding efficiency are the *Conversational High Compression* (CHC) profile and the *High Latency Profile* (HLP). The CHC profile includes most of the optional features (Annexes D, F, I, J, T, and U) that provide enhanced coding efficiency for low-delay applications. The HLP profile adds the support of B pictures (as defined in Annex O) to the coding efficiency tools of the CHC profile and is targeted for applications that allow a higher coding delay.

The 2nd generation video coding standard AVC/H.264 was developed in the period 1999-2009 from version 1 to 16 [6]. One of the most improvement is increased flexibility for **inter coding**. For the purpose of motion-compensated prediction, a macroblock can be partitioned into square and rectangular block shapes. H.264/MPEG-4 AVC also supports multiple reference pictures. Similarly to Annex U of H.263, motion vectors are associated with a reference picture index for specifying the employed reference picture. The motion vectors are transmitted using quarter-sample precision relative to the luma sampling grid. Luma prediction values at half-sample locations are generated using a 6-tap interpolation filter and prediction values at quarter-sample locations are obtained by averaging two values at integer- and half-sample positions. Weighted prediction can be applied using a scaling and offset of the prediction signal. In general, motion vectors are predicted by the component-wise median of the motion vectors of three neighboring previously decoded blocks. In contrast to prior coding standards, the concept of B pictures is generalized and the picture coding type is decoupled from the coding order and the usage as a reference picture. Instead of I, P, and B pictures, the standard actually specifies I, P, and B *slices*. A picture can contain slices of different types and a picture can be used as a reference for inter prediction of subsequent pictures independently of its slice coding types. This generalization allowed the usage of prediction structures such as hierarchical B pictures [13]. AVC/H.264 also modified design for **intra coding**. While in previous standards some of the DCT coefficients can be predicted from neighboring intra blocks, the intra prediction in AVC/H.264 is done in the spatial domain by referring to neighboring samples of previously decoded blocks. For **transform coding**, AVC/H.264 specifies a 4×4 and an 8×8 transform. While chroma blocks are always coded using the 4×4 transform, the transform size for the luma component can be selected on a macroblock basis. For intra MBs, the transform size is coupled to the employed intra prediction block size. An additional 2×2 Hadamard transform is applied to the four DC coefficients of each chroma component. For the intra 16×16 mode, a similar second-level Hadamard transform is also applied to the 4×4 DC coefficients of the luma signal. In contrast to previous standards, the inverse transforms are specified by exact integer operations, so that, in error-free environments, the reconstructed pictures in the encoder and decoder are always exactly the same. The transform coefficients are represented using a uniform reconstruction quantizer, i.e., without the extra-wide dead-zone that is found in older standards. Similar to H.262/MPEG-2 Video, AVC/H.264 also supports the usage of quantization weighting matrices. The transform coefficient levels of a block are generally scanned in a zig-zag fashion. For entropy coding of all macroblock syntax elements, AVC/H.264 specifies two methods: CAVLC and CABAC. Context-adaptive variable-length coding (CAVLC), uses a single codeword set for all syntax elements except the transform coefficient levels. The approach for coding the transform coefficients basically uses the concept of run-level coding as in prior standards. However, the efficiency is improved by switching between VLC tables depending on the values of previously transmitted syntax elements. Context-adaptive binary arithmetic coding (CABAC) improve coding efficiency relative to CAVLC. The statistics of previously coded symbols are used for estimating conditional probabilities for binary symbols, which are trans-mitted using arithmetic coding. Inter-symbol dependencies are exploited by switching between several estimated probability models based on previously decoded symbols in neighboring blocks. Similar to Annex J of H.263, AVC/H.264 includes a deblocking filter inside the motion compensation loop. The strength of the filtering is adaptively controlled by the values of several syntax elements. The High Profile of AVC/H.264 includes all tools that contribute to the coding efficiency for 8-bit-per-sample video in 4:2:0 format [4].

The 3rd generation video coding standard HEVC/H.265 was developed in the period 2010-2013 as version 1 [2]. HEVC retains the basic hybrid coding architecture (Fig.1a,b) of prior video coding standard AVC/H.264. A significant difference lies in the use of a more adaptive quadtree structure based on a *coding tree unit* (CTU) instead of *macroblock*. In principle, the **quadtree coding structure** is described by means of *blocks* and *units*. A unit encapsulates one luma and corresponding chroma blocks with syntax needed to code these. Consequently, a CTU includes *coding tree blocks* (CTBs) and syntax specifying coding data and further levels with *coding blocks* (CBs). Each CU incorporates more *prediction units* (PU), and *transform units* (TU). Similarly, each CB is split into *prediction blocks* (PB) and *transform blocks* (TB) (Fig.1d). This variable size adaptive approach is particularly suited to larger resolutions, such as HD formats. **Intra-prediction** in HEVC is quite similar to AVC/H.264. Samples are predicted from reconstructed samples of neighboring blocks. The mode categories remain identical: *horizontal/vertical*, *directional, DC* and *plane*. A significant change comes from the introduction of larger block size, where intra-prediction using one of the 35 modes may performed four blocks of size up to 32x32 samples. The increased number of predictive modes (35 in HEVC vs. 8 in AVC/H.264) requires efficient mode selection heuristics, to reduce complexity of encoders. As in AVC/H.264, uniform reconstruction quantization (URQ) is used in HEVC, with quantization scaling matrices supported for the various transform block sizes. **Inter-prediction** based on motion compensation (MCP), is conceptually very simple in HEVC, but comes with some overhead compared to AVC/H.264. HEVC supports quarter-sample precision *motion vectors* in MCP. The luma prediction signal for all fractional-sample location is generated by separable 7- or 8-tap filters. For chroma, 4-tap interpolation filters are applied. HEVC also supports multiple reference pictures, while the concepts of I, P, and B *slices* are basically unchanged from AVC/H.264. The coding of motion parameters has been substantially improved compared to prior standards. HEVC supports a so-called **merge mode**, in which no motion parameters are coded. Instead, a candidate list of motion parameters is derived for the corresponding PU. Merge mode sets all motion parameters of an inter-predicted block equal to the parameters of a merge candidate. While AVC/H.264 downsamples motion vectors to the 8x8 levels, HEVC further reduces memory requirements by keeping a single motion vector per 16x16 blocks. Also, HEVC offers more ways to split a picture into motion compensated partition patterns. Besides a *deblocking filter*, the HEVC design includes a *sample-adaptive offset* (SAO) operation inside the motion compensation loop. SAO classifies the reconstructed samples into different categories, depending on sample amplitude or edge characteristics and reduces the distortion by adding a separate offset for each class of samples. HEVC includes three components that enable some degree of high-level parallelism: *slices*, *tiles* and *wavefronts* (Fig.1c). High level parallelism refers to the ability to simultaneously process multiple regions of a single picture. Support for such parallelism may be advantageous to both encoders and decoders where multiple identical processing cores may be used in parallel.

In summary, HEVC provides the following new features (Table 1):

- quad-tree partitioning for prediction and transform with more and larger block sizes,
- parallel processing with tiles and wavefronts,
- ultra-low delay processing with dependent slices,
- Inter-picture prediction block merging,
- advanced motion vector prediction (AMVP),
- high-throughput transform coefficient coding,
- transform skip mode for screen content coding,
- sample adaptive offset in-loop filtering (SAO).

Table 1. Coding tools of HEVC MP vs. AVC HP.

	HEVC	**AVC**
Block size	Tree structure 8x8, 16x16, 32x32, 64x64 Square, sym./asym. rect.	Macroblock 16x16 (4x4) Square, sym. rect.
Transforms	Integer-DCT (4x4, 8x8, 16x16, 32x32) Integer-DST (4x4 Intra)	Integer-DCT (4x4, 8x8) Hadamard (2x2, 4x4)
Intra-prediction	up to 33 angular modes (+DC+planar mode)	up to 9 modes
Motion prediction Motion-copy mode MV precision	Advanced MV prediction (spatial+temp. co-located) Merge, Skip ¼ pixel 7/8-tap	Spatial median +temp. co-located Direct, Skip ½ pixel 6-tap + ¼ pixel bilinear
In-loop filtering	deblocking, SAO	deblocking
Quantization Entropy coding	URQ CABAC	URQ CAVLC, CABAC

2.1 LowDelay configuration of encoders

The HEVC/H.265, AVC/H.264, and H.263 supports low-delay coding structures that usually provide an improved coding efficiency. Since interactive applications require a low coding delay, all pictures were coded in display order, where only the first picture is coded as an intra picture and all subsequent pictures are temporally predicted only from reference pictures in the past in display order. Dyadic low-delay hierarchical prediction structures with groups of 4 pictures are selected [13]. The same 4 previously coded pictures are used as active reference pictures. All pictures are coded as a single slice.

For the low-delay coding conditions, only the first picture in a video sequence shall be encoded as IDR (*Instantaneous Decoding Refresh*) picture. In mandatory low-delay test condition, the other successive pictures shall be encoded as Generalized P and B-picture (GPB). Reference picture list combination *RefPicList0* and *RefPicList1* is used for management and entropy coding of reference picture index. Graphical presentation of LowDelay configuration is shown in Figure 1b. The number associated with each picture represents encoding order. QP of each inter coded picture shall be derived by adding offset to QP of Intra coded picture depending on temporal layer [2].

2.2 Encoder control

The purpose of the software implementation of HEVC Test Model (HM) encoder is to provide a *common reference* implementation of an HEVC encoder that is useful for evaluating technologies and for independent encoder/decoder development [3]. Compared to the AVC/H.264 JM reference software, the HM provides for fewer parameters to configure the encoder. The HM decoder is an example implementation aimed at *correctness*, *completeness* and *readability*. All reference implementation of encoders use the same strategies for mode decision, motion estimation, and quantization.

The task of an encoder control is to determine the values of the syntax elements, and thus the bitstream, for a given input sequence in a way that the distortion between the input sequence and its reconstruction is minimized subject to constraints for the average and maximum bit rate (*Operational ORD points*). The overall minimization problem is split into a series of smaller minimization problems by partly neglecting spatial and temporal interdependencies between coding decisions p

$$\min D_k(p) \text{ subject to } R_k(p) \leq R_c \qquad (1)$$

The used *distortion measures D* are defined for the sum of absolute differences (SAD) and for the sum of squared differences (SSD) between original and reconstructed samples. Except for motion estimation, SSD was used as the distortion measure for all coding decisions. Hence, all encoders are basically optimized with respect to the mean squared error (MSE) or peak signal-to-noise ratio (PSNR). The constrained minimization problem in (1) can be reformulated as an unconstrained minimization

$$\min D_k(p) + \lambda \cdot R_k(p) \qquad (2)$$

The concept of *Lagrangian multiplier* encoder control is applied for mode decision [14], motion estimation ME [13], and quantization URQ [15]. The *Lagrange m*ultiplier is set according to

$$\lambda = \alpha \cdot Q^2 \qquad (3)$$

where Q denotes the **quantization step size**, which is controlled by the **quantization parameter** *QP*. Given the quantization parameter for intra pictures, the quantization parameters for all other pictures and the factors are set using a deterministic approach.

3 Results

For comparing the coding efficiency of HEVC with AVC/H.264 and H.263 video coding standards, coding experiments for the scenario of interactive applications where performed. For HEVC, the described encoder control is the same as the one implemented in the HM 8.0 reference software implementation [17]. For the AVC/H.264 and H.263 standards, the described encoder control was integrated into encoder implementations [18, 19]. All encoders use the same strategies for mode decision, motion estimation, and quantization. The quantization parameter and the *Lagrange multiplier* are held *constant* for all macroblocks or coding units of a video picture. For HEVC, all coding tools specified in draft HEVC/H.265 Main Profile, AVC/H.264 High Profile (HP) and H.263 Conversational High Compression (CHC) profile are enabled.

3.1 PSNR/MOS performance

The three test sequences (class E') with typical video conferencing content was selected in experiments (*Vidyo1/2/3* 1280x720 60fps x10s). Each test sequence was coded at 12 different bitrates (Table 2). For HEVC MP and AVC/H.264 HP, the quantization parameter QP for intra pictures was varied in the range 20-42. For H.263 CHC the quantization parameters for intra pictures were chosen in a way that the resulting quantization step sizes are approximately the same as for HEVC and AVC.

Table 2. Operational ORD points of HEVC, AVC and H.263 for 12 different quantization parameters and bitrates of test sequence *Vidyo1*.

Quantization Parameter QP Bitrate [Mbps] PSNR [dB]	HEVC	AVC	H.263
QP$_1$	20	20	2
QP$_2$	22	22	3
QP$_3$ BR PSNR$_Y$ PSNR$_U$ PSNR$_V$	24 1.292 42.49 46.17 47.02	24 1.709 42.04 46.00 46.81	4
QP$_4$ BR PSNR$_Y$ PSNR$_U$ PSNR$_V$	26 0.876 41.57 45.59 46.31	26	5
QP$_5$ BR PSNR$_Y$ PSNR$_U$ PSNR$_V$	28	28 0.849 40.05 44.74 45.22	6 1.357 39.70 44.88 45.30
QP$_6$ BR PSNR$_Y$ PSNR$_U$ PSNR$_V$	30 0.466 39.66 44.67 45.25	30	8
QP$_7$ BR PSNR$_Y$ PSNR$_U$ PSNR$_V$	32 0.351 38.61 43.92 44.37	32 0.499 37.87 43.69 43.82	10 0.846 37.40 43.54 43.60
QP$_8$ BR PSNR$_Y$ PSNR$_U$ PSNR$_V$	34 0.272 37.50 43.44 43.70	34 0.392 36.77 43.20 43.31	13
QP$_8$	36	36	16
QP$_{10}$ BR PSNR$_Y$ PSNR$_U$ PSNR$_V$	38	38 0.251 34.49 42.13 41.87	20 0.533 34.00 42.75 42.81
QP$_{11}$	40	40	25
QP$_{12}$	42	42	31

The *Rate-Distortion* (RD) curves of the combined luma and chroma components are used for bitrates 0.256, 0.384, 0.512, 0.850, and 1.500Mbps (Fig. 2). The combined PSNR$_{YUV}$ is first calculated as the weighted sum of the PSNR per picture of the individual components (PSNR$_Y$, PSNR$_U$, PSNR$_V$) to obtain

$$PSNR_{YUV} = (6 \cdot PSNR_Y + PSNR_U + PSNR_V)/8 \quad (4)$$

where individual components are computed as

$$PSNR = 10 \log_{10} (2^B-1)^2/MSE \quad (5)$$

Here, *B*=8 is the number of bits per sample of the video signal to be coded, while the mean square error MSE is the aim of the squared differences SSD

divided by the number of samples in the signal. The PSNR measurements per video sequence are computed by averaging the per-picture measurements.

Bitrate savings of three typical HD720 tested sequences are shown in Table 3. HEVC provides significant gains in term of average compression efficiency of 59.35% and 148.76% relative to AVC and H.263, respectively. As can be seen, the HEVC coding efficiency gains for the lower bitrate range are generally higher than the average results.

Table 3. Coding gain of HEVC LD over AVC HP / H.263 CHC for typical bitrates: BitRate reduction [Mbps] based on ORD interpolation and PSNR=*const*.

HD720 *Vidyo1*	HEVC	PSNR=41.25dB BR=0.512Mbps	PSNR=42.59dB BR=0.850Mbps	PSNR=43.79dB BR=1.500Mbps
	AVC	BRr=0.329 64.26%	BRr=0.508 59.76%	BRr=0.910 60.73%
	H.263	BRr=0.928 181.25%	BRr=1.275 150%	BRr=1.895 126.33%
HD720 *Vidyo2*	HEVC	PSNR=40.65dB BR=0.512Mbps	PSNR=42.04dB BR=0.850Mbps	PSNR=43.29dB BR=1.500Mbps
	AVC	BRr=0.289 56.44%	BRr=0.478 56.24%	BRr=0.860 57.33%
	H.263	BRr=1.142 223.00%	BRr=1.658 200.00%	BRr=2.318 154.53%
HD720 *Vidyo3*	HEVC	PSNR=41.19dB BR=0.512Mbps	PSNR=42.42dB BR=0.850Mbps	PSNR=43.60dB BR=1.500Mbps
	AVC	BRr=0.316 61.72%	BRr=0.495 58.24%	BRr=0.900 60%
	H.263	BRr=1.054 205.86%	BRr=1.398 164.47%	BRr=2.016 134.40%

PSNR coding gain of three typical HD720 tested sequences are shown in Table 4. HEVC provides significant gains in term of average coding efficiency of 3.16% and 10.08% relative to AVC and H.263, respectively. As can be seen, the HEVC coding gains for the lower bitrate range are generally higher than the average results.

In Figures 3 and 4 are depicted the variability of PSNR/BR vs. Frames (time) graphs of the three test signals encoded by HEVC/AVC/H.263 profiles. Across all the encoded signals, the encoding parameters have remained constant in order to be possible the comparison between the various encoders and the different signal types. As it can be deduced by these figures that the *encoding efficiency* of the HEVC is similar to the AVC, while it doubles the *compression efficiency* of the bitstream. Next, although the variability of PSNR/BR vs. Frames (time) appear similar, H.263 encoding performance is smoother in comparison to the HEVC that have greater variance.

Table 4. Coding gain of HEVC LD over AVC HP / H.263 CHC for typical bitrates: $PSNR_{Y\ CodingGain}$ [dB] based on ORD interpolation and BR=*const.*

		BR=0.512Mbps PSNR=41.25dB	BR=0.850Mbps PSNR=42.59dB	BR=1.500Mbps PSNR=43.79dB
HD720 *Vidyo1*	HEVC			
	AVC	$PSNR_{CG}$=1.82 4.60%	$PSNR_{CG}$=1.31 3.18%	$PSNR_{CG}$=0.95 2.21%
	H.263	$PSNR_{CG}$=5.33 13.86%	$PSNR_{CG}$=3.62 9.00%	$PSNR_{CG}$=2.40 5.80%
HD720 *Vidyo2*	HEVC	BR=0.512Mbps PSNR=40.65dB	BR=0.850Mbps PSNR=42.04dB	BR=1.500Mbps PSNR=43.29dB
	AVC	$PSNR_{CG}$=1.61 4.12%	$PSNR_{CG}$=1.20 2.94%	$PSNR_{CG}$=0.97 2.29%
	H.263	$PSNR_{CG}$=9.71 31.39%	$PSNR_{CG}$=4.34 11.51%	$PSNR_{CG}$=2.96 7.35%
HD720 *Vidyo3*	HEVC	BR=0.512Mbps PSNR=41.19dB	BR=0.850Mbps PSNR=42.42dB	BR=1.500Mbps PSNR=43.60dB
	AVC	$PSNR_{CG}$=1.61 4.07%	$PSNR_{CG}$=1.18 2.85%	$PSNR_{CG}$=0.92 2.16%
	H.263	$PSNR_{CG}$=5.89 16.67%	$PSNR_{CG}$=3.93 10.20%	$PSNR_{CG}$=2.55 6.22%

In subjective quality performance test, a group of subjects is asked to watch a set of video sequences (Fig. 2), and to rate their visual quality by using a particular rating scale. The scores assigned by the observers to each test stimulus are averaged in order to obtain a mean opinion score (MOS). The subjective assessment are set up following ITU-R Rec. BT.500 [20]. Double stimulus impairment scale (DSIS) is used for the evaluation of the quality. A quality rating scale made of 11 levels is adopted, ranging from 0 (*lowest quality*) to 10 (*highest quality*). The structure of the *Basic Test Cell* of the DSIS method consists of two consecutive presentations of the sequence under test. The MOS values were computed from the votes provided by the subjects for each test point [21].

Based on subjective results shown in Table 5 for typical videoconferencing application, the overall average bitrate reduction was 66.33%. As with the objective measures, the savings are largest at the lowest bitrates.

Table 5. BitRate reduction of HEVC CfP (best performing proponent 03) vs. AVC HP (Anchor B) based on subjective MOS performance for typical video conferencing bitrates [21].

Test sequence	BRr	average BRr
HD720 *Vidyo1*	59%	
HD720 *Vidyo2*	69%	66.33%
HD720 *Vidyo3*	71%	

3.2 Video codec complexity

The encoding and decoding times for the representative HD720 sequences (60fps x 10s) from the JCT-VC common test conditions, each of which is 10 seconds long, are shown in Table 6. Results are limited to the one quantization parameter (QP) values defined in the test conditions [22]. Times are recorded in 10s of seconds such as to illustrate the ratio to real-time operation. The HEVC encoding time exceed 1000 times real-time (Table 6). The decoding time exceed 4 times real-time on an *Ultrabook* x86 Core i5 2/4@1.7GHz 4GB RAM.

Table 6. Complexity of HEVC HM8 LD vs. AVC HP vs. H.263 CHC based on encoding/decoding time [10s] .

HD720 *Vidyo1* **60fps** x10s=600frames BR=663.552Mbps	HEVC LD	QP=24 BR=1.294Mbps TimeENC=1959 TimeDEC=25 fps
	AVC HP	QP=24 BR=1.708Mbps TimeENC=562 TimeDEC=12 fps
	H.263 CHC	QP=6 BR=1.357Mbps TimeDEC=30 fps
HD720 *Vidyo2* **60fps** x10s=600frames BR=663.552Mbps	HEVC LD	QP=24 BR=1.658Mbps TimeENC=1902 TimeDEC=25 fps
	AVC HP	QP=26 BR=1.421Mbps TimeENC=541 TimeDEC=12fps
	H.263 CHC	QP8= BR=1.225Mbps TimeDEC=30 fps
HD720 *Vidyo3* **60fps** x10s=600frames BR=663.552Mbps	HEVC LD	QP=24 BR=1.383Mbps TimeENC=2008 TimeDEC=25 fps
	AVC HP	QP=26 BR=1.188Mbps TimeENC=665 TimeDEC=12 fps
	H.263 CHC	QP=6 BR=1.438Mbps TimeDEC=30 fps

4 Conclusion

The third generation HEVC/H.265 video coding standard is designed to achieve multiple goals, including coding efficiency, easy of transport system integration, as well as parallel processing architectures. The complexity of HEVC decoder is not significantly different from that of AVC/H.264 decoders, making HEVC decoding in software practical on current hardware. We studied HEVC codec efficiency in real-time conversational application under unified testing approach.

The HEVC video codec provides significant increased coding efficiency compared to AVC/H.264 and H.263 standards under the low-delay encoding constraints typical for videoconferencing applications. The results of tests for selected video sequences indicate that

HEVC/H.265 encoders can achieve equivalent objective video quality as encoders that conform to AVC/H.264 when using approximately 60% less bit rate on average. Bitrate reduction is based on PSNR measure, while complexity is based on decoding time on personal *Ultrabook* computer in conversational application.

Statistical multiplexing of multiple encoded video streams, digital video transcoding and error robustness in multimedia communication systems is expected to be an active research area in years to come.

References:

[1] K.R.Rao, Z.S.Bojkovic, D.A.Milovanovic, *Introduction to multimedia communications: applications,middleware,networking*, Wiley, 2005.

[2] JCT-VC L1003 ITU-T SG 16 WP3 and ISO/IEC JTC 1/SC 29/WG 11, *High Efficiency Video Coding (HEVC) text specification draft 10*, 12th Meeting: Geneva, CH, 14–23 Jan. 2013.

[3] JCT-VC J1002 ITU-T SG 16 WP3 and ISO/IEC JTC 1/SC 29/WG 11, *HM8: High Efficiency Video Coding (HEVC) Test Model 8 Encoder description*, 10th Meeting: Stockholm, SE, 11–20 July 2012.

[4] G.J.Sullivan, J-R.Ohm, W-J.Han, T.Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard", *IEEE Trans. CSVT*, vol.22, no.12, Dec. 2012, pp.1649-1668.

[5] ITU-T and ISO/IEC JTC1, *Advanced Video Coding for generic audio-visual services*, ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC), version 1: 2003, version 2: 2004, versions 3, 4: 2005, versions 5, 6: 2006, version, s 7, 8: 2007, versions 9, 10, 11: 2009, versions 12, 13: 2010, versions 14, 15: 2011, version 16: 2012.

[6] D.Milovanović, Z.Bojković, "Advanced video coding standard H.264/AVC: performance vs. complexity", *WSEAS Trans. on Communications*, Issue 1, Volume 4, 2005, pp.1-6.

[7] S.-G.Cho, Z.Bojković, D.Milovanović, J.Lee, J.-J.Hwang, "Image quality evaluation: JPEG 2000 versus Intra-only H.264/AVC HighProfile", *WSEAS Int.Conf. on Communications*, Oct. 2006, Bucharest, Romania.

[8] T.Moorthy, P.P.Chen, A.Ye, "A scalable architecture for H.264/AVC variable block size motion estimation on FPGAs", *WSEAS Trans. Signal proc*, Volume 7, January 2011, pp. 23-33.

[9] F. Neri, "Cooperative evolutive concept learning: an empirical study", *WSEAS Trans. on Information Science and Applications*,WSEAS Press (USA), issue 5, vol. 2, 2005, pp. 559-563.

[10] J-R.Ohm, G.J.Sullivan, H.Schwarz, T.K.Tan, T.Wiegand, "Comparison of the coding efficiency of video coding standards - including High Efficiency Video Coding (HEVC)", *IEEE Trans. CSVT*, vol.22, no.12, 2012, pp.1669-1684.

[11] ITU-T, *Video coding for low bitrate communication*, ITU-T Rec. H.263, version 1, 1995, version 2, 1998, version 3, 2000.

[12] ITU-T and ISO/IEC JTC 1, *Generic coding of moving pictures and associated audio information* – Part 2: Video, ITU-T Rec. H.262 ISO/IEC 13818-2 (MPEG-2), version 1: 1994.

[13] H. Schwarz, D. Marpe, T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. CSVT*, vol. 17, Sep. 2007, pp. 1103–1120.

[14] G.J.Sullivan, T.Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Proc. Magazine*, vol.15, 1998, pp.74-90.

[15] T.Wiegand, M.Lightstone, D.Mukherjee, T.G.Campbell, S.K.Mitra, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard," *IEEE Trans. CSVT*, vol. 6, Apr. 1996, pp. 182–190.

[16] E.-H.Yang, X.Yu, "Rate distortion optimization for H.264 interframe coding: A general framework and algorithms," *IEEE Trans. Image Processing*, vol.16, 2007, pp.1774–1784.

[17] JCT-VC of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, *HM-8.0 reference software*, https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-8.0/

[18] JVT-AF013, *Joint Scalable Video Model software*, V9.19, *Joint Video Team*, Nov.2009.

[19] T.Wiegand, H.Schwarz, A.Joch, F.Kossentini, G.J.Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. CSVT*, vol.13, no.7, July 2003, pp. 688–703.

[20] Z.Milicevic, Z.Bojkovic, K.R.Rao, "An approach to interactive multimedia systems through subjective video quality assessment in H.264/AVC standard", *WSEAS Trans. Systems,* Issue 8, Vol. 11, August 2012, pp.305- 314.

[21] F.DeSimone, L.Goldmann, J-S.Lee, T.Ebrahimi, "Towards high efficiency video coding: Subjective evaluation of potential coding technologies", *Journal of Visual commun. and Image representation*, vol.22, issue8, Nov. 2011, pp.734-748.

[22] F.Bossen, B.Bross, K.Suhring, D.Flynn, "HEVC complexity and implementation analysis", *IEEE Trans. CSVT*, vol.22, no.12, Dec. 2012, pp.1685-1696.
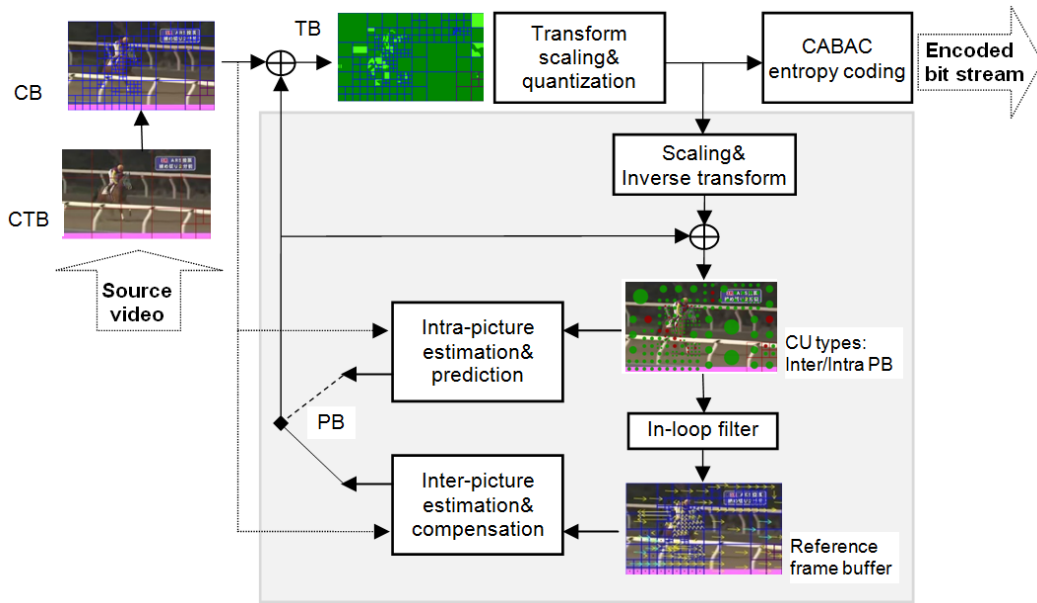
a) HEVC block-based hybrid video encoder.
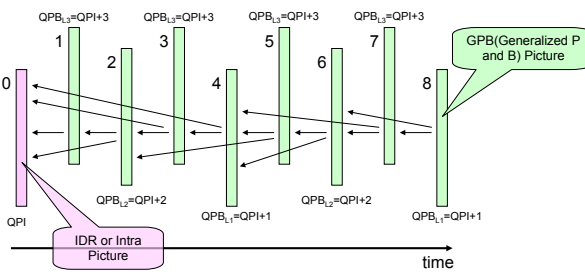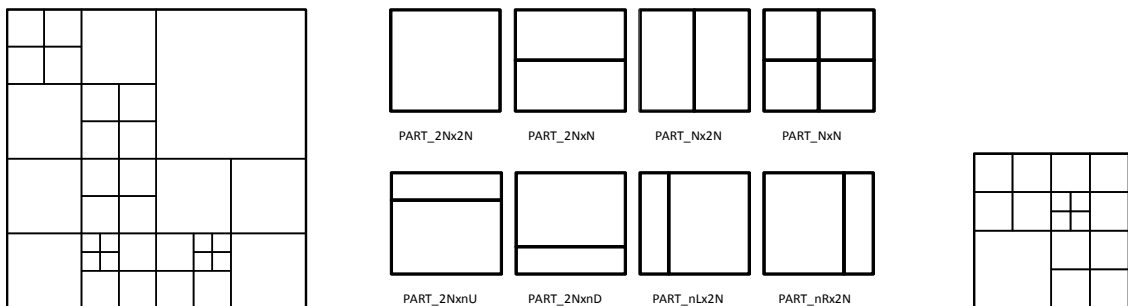

b) HEVC motion-compensated prediction in LowDelay configuration [2].
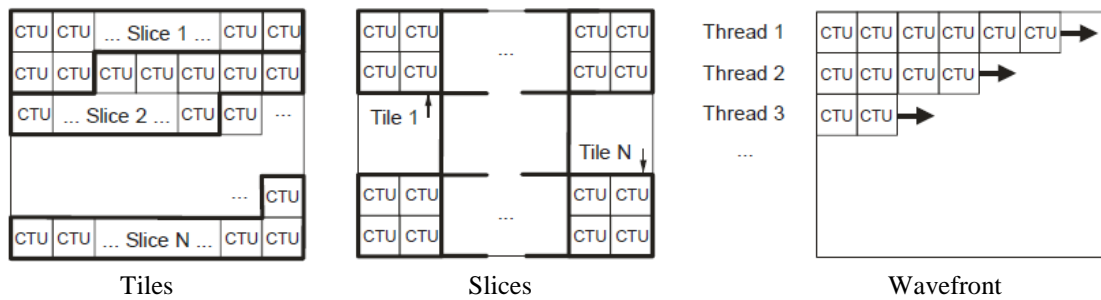
c)　　　　　Tiles　　　　　　　　　　Slices　　　　　　　　　Wavefront

d)　　　Coding Blocks (CB)　　　　　PredictionBlocks (PB)　　　　TransformBlocks (TB)

Fig. 1   a) HEVC block-based hybrid transform-entropy video encoder, b) motion-compensated prediction, c) picture tiles and slices, and wavefront parallel processing, d) adaptive picture partitioning example of coding quadtree CTU into CU (CB), partition modes for PU (PB), and transform quadtree within CU (TB).
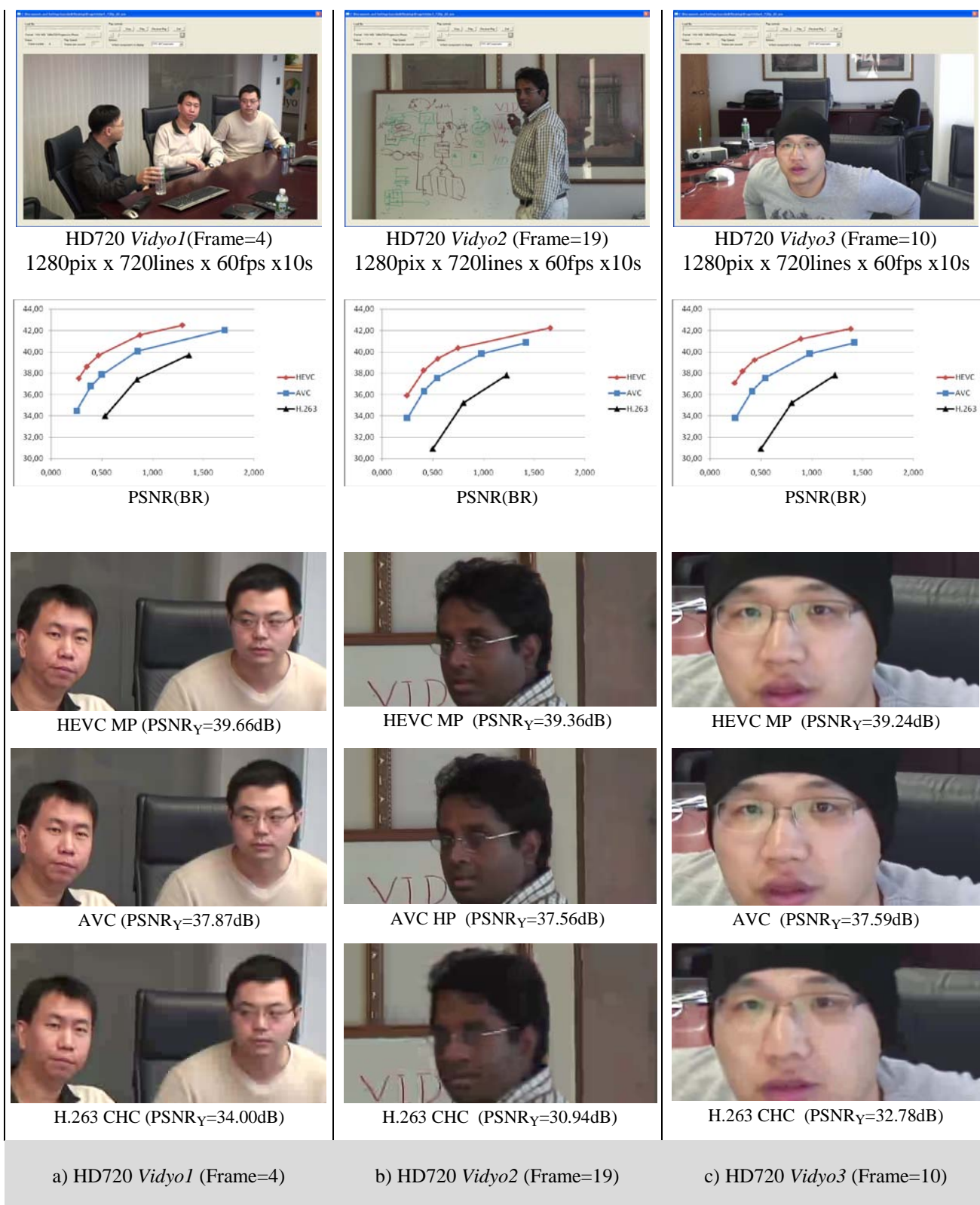
Fig. 2 Coding performance $PSNR_Y$(BR) of HEVC MP LD vs. AVC HP vs. H.263 CHC
and HEVC subjective quality gain of the sample frames
encoded at bitrate BR~0.512Mbps (QP_{HEVC}=30, QP_{AVC}=32, $QP_{H.263}$=20/31/25)
a) HD720 *Vidyo1*, b) HD720 *Vidyo2* , d) HD720 *Vidyo3* test sequence.

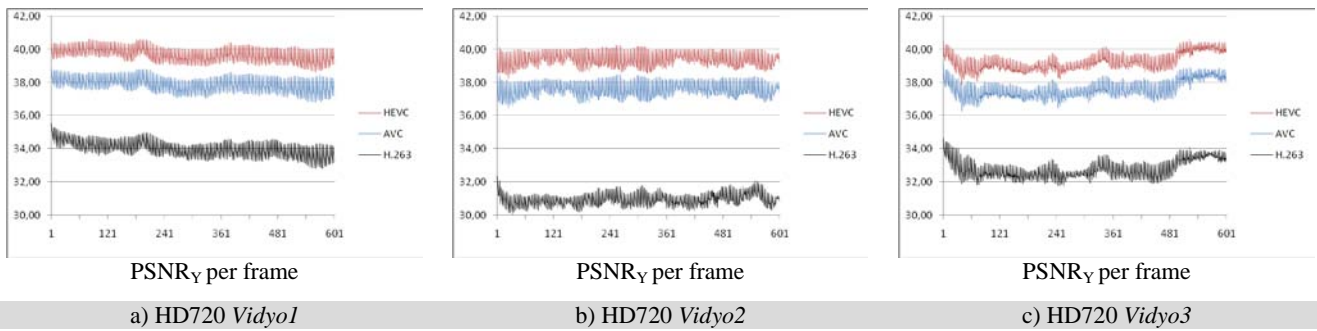| a) HD720 *Vidyo1* | b) HD720 *Vidyo2* | c) HD720 *Vidyo3* |

Fig. 3 Variability PSNR$_Y$ per frame (time) for BR=*const*
(BR~0.512Mbps: QP$_{HEVC}$=30, QP$_{AVC}$=32, QP$_{H.263}$=20/31/25):
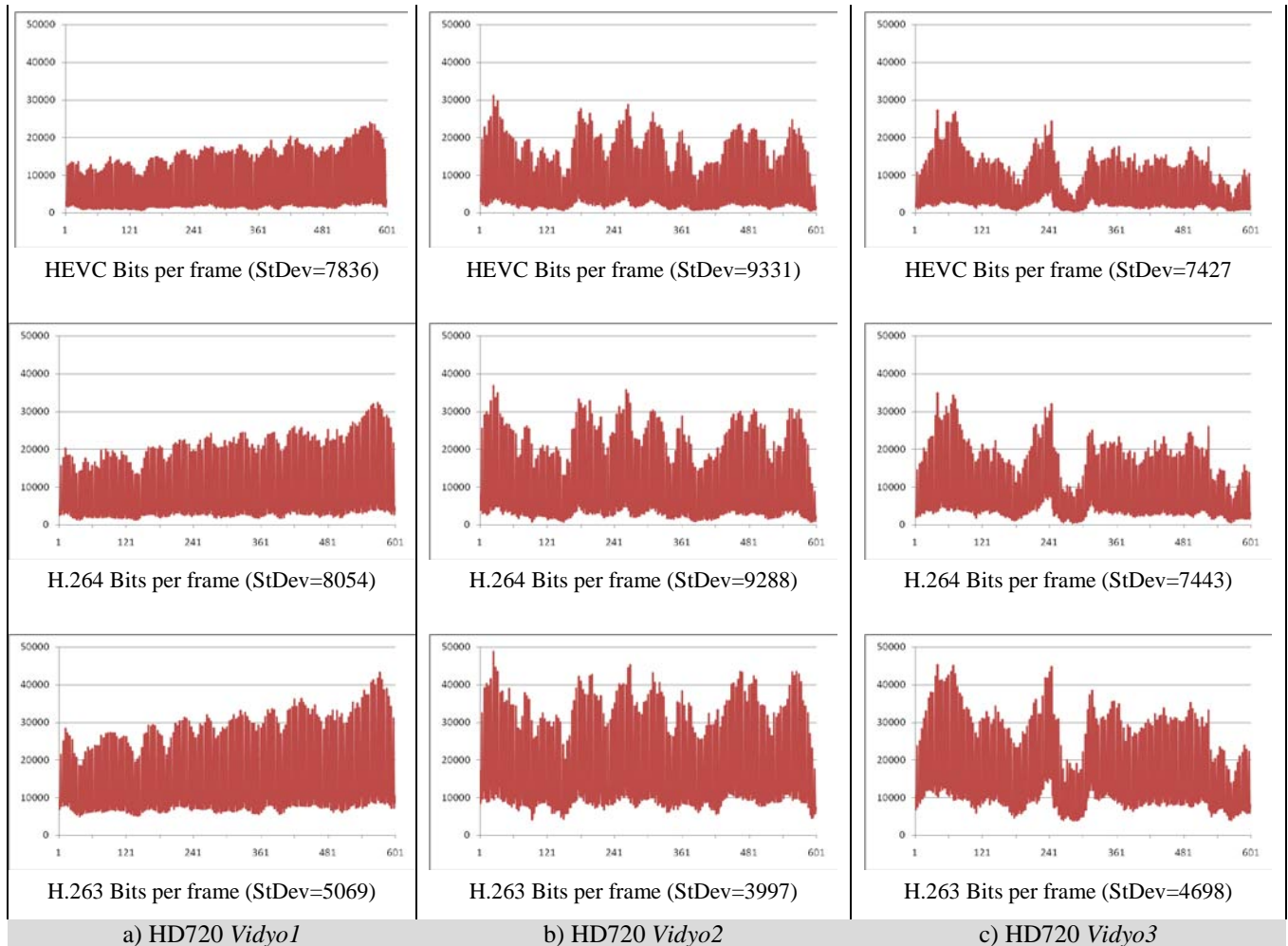a) HD720 *Vidyo1*, b) HD720 *Vidyo2*, c) HD720 *Vidyo3*.

HEVC Bits per frame (StDev=7836)	HEVC Bits per frame (StDev=9331)	HEVC Bits per frame (StDev=7427
H.264 Bits per frame (StDev=8054)	H.264 Bits per frame (StDev=9288)	H.264 Bits per frame (StDev=7443)
H.263 Bits per frame (StDev=5069)	H.263 Bits per frame (StDev=3997)	H.263 Bits per frame (StDev=4698)
a) HD720 *Vidyo1*	b) HD720 *Vidyo2*	c) HD720 *Vidyo3*

Fig. 4 Variability bitrate BR per frame (time) for PSNR$_Y$=*const* (QP$_{HEVC}$=32, QP$_{AVC}$=32, QP$_{H.263}$=10):
a) HD720 *Vidyo1*, b) HD720 *Vidyo2*, c) HD720 *Vidyo3*.