Compression Performance and Video Quality Comparison of HEVC and AVC

Harilaos Koumaras
NCSR Demokritos
Institute of Informatics and Telecommunications
Aghia Paraskevi, Greece
koumaras@iit.demokritos.gr

Michail-Alexandros Kourtis, Charalampos Skianis
University of the Aegean
Dept. of Information & Communication Systems Eng.
Karlovasi, Samos
{akiskourtis, cskianis}@aegean.gr

Abstract: **The main purpose of the paper is to perform a quantitative and qualitative performance evaluation of the emerging standard H.265/HEVC in comparison to its predecessor H.264/AVC. The paper focuses on both the encoding and compression efficiency by measuring the PSNR values, SSIM scores and bit rate values of reference and non-reference signals. For the needs of the paper, the HEVC and AVC reference encoders have been used. Based on the experimental results, it is deduced that the HEVC meets its primary objective, which is to double the compression efficiency of the bit stream without significant degradation of the encoded video quality. Finally, interesting findings are also reported for the very low bit rate area of the encoding rate for low spatial resolutions.**

Keywords—HEVC; AVC; Benchmark; PSNR; SSIM;

I. INTRODUCTION

The rise of the video services and applications has brought the modern world closer to digital videos than ever before, creating new demands for high performance video compression standards [1], [2] in conjunction with efficient and resilient transmission systems [3]. This trend has created the need for high efficiency video coding methods and encourages various standardized efforts that aim to derive new and more efficient encoding standards in terms of both performance (i.e., Data Compression) and video quality (i.e., Quality of Experience). In this framework, ITU-T/VCEG and ISO-IEC/MPEG have recently formed the joint collaborative team on Video Coding (JCT-VC) with scope to develop the next generation video coding standard called High Efficiency Video Coding (HEVC) or H.265, which is being developed as the successor to H.264/AVC.

The HEVC's main goal is to substantially improve coding efficiency compared to AVC High Profile in terms of video compression and performance efficiency. This improvement is supported by 50% bitrate reduction (i.e., compression efficiency), while the video quality level will remain constant, probably at the expense of increased computational complexity, which is expected to be three times higher. Thus, the new codec aims to satisfy the current demands for cost effective video encoding process in terms of better compression efficiency and video quality.

In this basis, the authors of the paper, following the development of the reference HEVC test model (HM), report on the encoding performance and compression efficiency of currently working version of HEVC in respect with its main objectives. Moreover, for comparison and benchmarking reasons of the emerging new code, a set of YUV video signals (reference and non-reference ones) used as input to both the HEVC HM v5.1 and AVC JM v17 codec. The specific set of test signals covers a variety of spatiotemporal activity, making the benchmarking methodology of this paper appropriate for testing the performance efficiency of the new encoder under different coding conditions. To this extent the PSNR and SSIM metrics were selected for comparing the performance of the two codecs.

Upon this introductory section, the rest of the paper is organized as follows: Section 2 introduces the main novel features of HEVC. Section 3 describes the performance and video quality metrics used in this paper. Section 4 describes the encoding process of AVC and HEVC signals. In section 5 the benchmarking between the two codecs under test is performed. Finally, Section 6 concludes the paper.

II. INTRODUCING HEVC FEATURES

The HEVC continues to implement the block-based hybrid video coding framework [4], with the exception of the increased macroblock size (up to 64*x*64) compared to AVC. However, three novel block concepts are introduced, namely: the Coding Unit (CU), the Prediction Unit (PU) and the Transform Unit (TU). This approach allows the proposed codec to be easily adapted at various content types, applications, or devices that have diverse capabilities.

A. Coding Unit (CU)

CU is the basic coding unit like the H.264/AVC's macroblock and sub-macroblock, however the main difference lies in the fact that CU can have various sizes, without distinction corresponding to its size, but is restricted to be square shaped. All processing except frame-based loop filtering is performed on a CU-basis, including intra/inter prediction, transform, quantization and entropy coding. For each encoding process two values are also defined: the largest coding unit (LCU) and the smallest coding unit (SCU). It is assumed that a picture consists of non-overlapped LCUs [5].

In the case of HEVC implementation, CU is expressed as a recursive quadtree representation (i.e., a tree data structure in which each internal node has exactly four children). Figure 1 shows an example where LCU size is 128 and the maximum hierarchical depth is 2. The recursive structure for each CU is represented by the respective split flags. When the split flag is set to zero, then the coding of CU is performed in the current depth. When the split flag is set to one, CU_d (CU of depth d and size 2N*x*2N) is split into 4 independent CU_{d+1} which have depth (d+1) and size N*x*N.

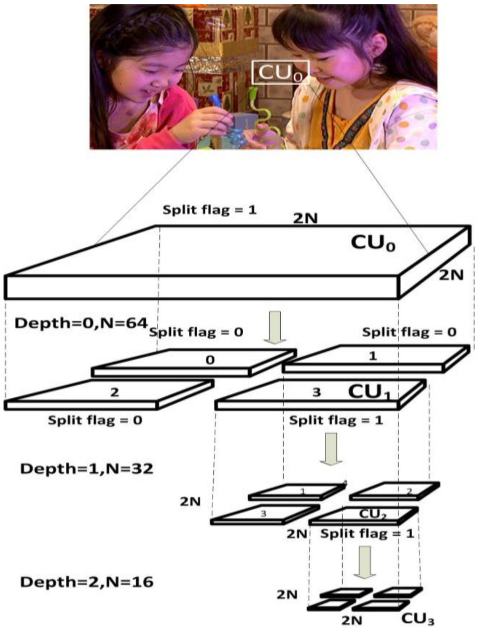


Fig. 1. Quadtree CU structure of HEVC.

In this case, CU_{d+1} is called a sub-CU of CU_d similar to a sub-macroblock in H.264/AVC. Depending on the maximum allowed depth (two in this case), this procedure with the flags continues recursively and each sub-$CU_{(d+1)}$ may further split to other four sub-CU of depth d+2, d+3... d+n. If the maximum allowed depth has been reached, then further splitting is not allowed and the process is terminated [5], [6].

B. *Prediction Unit (PU)*

Once the size of the Largest Coding Unit (LCU) and the hierarchical depth of CU are defined, then the leaf nodes CUs can be further split into PUs.

At the PU level, two different terms are introduced to specify the prediction method: PU type and PU splitting. Different PU splitting corresponds to different PU types, which consist of skip, intra and inter. The PU for intra has 2 different possible splittings: 2Nx2N (i.e. no split) and N*x*N (quarter split). The PU for inter has 8 different possible splittings: 4 symmetric splittings and 4 asymmetric splittings. A skipped PU can be 2N*x*2N (i.e. the whole CU is skipped).

C. *Transform Unit (TU)*

In addition to the CU and PU definitions, the transform unit (TU) is defined for transform and quantization purposes. In TU structure representation, residual quadtree structure is adopted, applying the same maximum quadtree depth for both luma and chroma components of each CU. It should be noted that the size of TU may be larger than that of the PU but not exceeding that of the CU.

TU have different splittings for low complexity (LC) and high efficiency (HE) configurations. For LC configuration, there are only two possible splittings from a CU to TUs, where the residual quadtree structure is restricted to two levels in LC configuration, while for the HE case the splitting from a CU to TUs is done recursively up to a maximum of three levels of the residual quadtree.

D. *Intra-Prediction in HEVC*

The current intra prediction technique in HEVC unifies two simplified directional intra-prediction methods: the Arbitrary Direction Intra and the Angular Intra Prediction. The unified intra prediction technique enables a lower-complexity method in which parallel processing can be supported, achieving better performance. More specifically, samples of already decoded adjacent PUs are used in order to define the type of the intra prediction method (i.e., horizontal, vertical or depending on the block size up to 28 angular directions) [7].

E. *Inter-Prediction in HEVC*

The inter prediction in HEVC uses the frames stored in a reference frame buffer, which allows multiple bi-direction frame reference. A reference picture index and a motion vector displacement are needed in order to select reference area. The merging of adjacent PUs is possible, by the motion vector, not necessarily of rectangular shape as their parent CUs. In order to achieve encoding efficiency, skip and direct modes similar to the AVC ones are defined, and motion vector derivation or a new scheme named motion vector competition is performed on adjacent PUs. Motion compensation is performed with a quarter-sample motion vector precision. At TU level (which commonly is not larger than the PU), an integer spatial transform (with range from 4*x*4 to 64*x*64) is used, similar in concept to the DCT transform. In addition a rotational transform can be used for block sizes larger than 8*x*8, and apply only to lower frequency components. In AVC scaling, quantization and scanning of transform are performed in a similar way.

At CU level, an Adaptive Loop Filter (ALF) can be applied prior to copying the frame into the reference picture buffer. This is a FIR filter whose main purpose is to minimize distortion relative to the original picture, and its filter coefficients which are encoded at slice level. Additionally a deblocking filter is operated within the prediction loop (similar to the AVC deblocking filter design). After applying these two filters the display output is registered to the buffer.

III. PERFORMANCE METRICS

For the experimental needs of the paper, the following objective metrics were used for evaluating purposes, namely the PSNR and SSIM, which are analysed hereby briefly..

A. *PSNR*

The PSNR performance metric [8] was used in order to quantify the performance enhancement of HEVC in comparison to AVC. The PSNR metric is defined in the following function:

$$PSNR = 10 \times \log_{10} \frac{MaxErr^2 \times w \times h}{\sum_{i=0,j=0}^{w,h} (x_{i,j} - y_{i,j})^2} \qquad (1)$$

The PSNR metric is mostly used as a measure to assess the noise introduced during the encoding process. PSNR is defined via the mean squared error for two w(idth) *x* h(eight) monochrome images $x_{i,j}$ and $y_{i,j}$ where one of the images is considered a noisy approximation of the other, with *MaxErr* being the maximum possible absolute value of color components difference.

B. SSIM

The SSIM is a Full Reference (FR) objective video quality metric, which measures the structural similarity between two images/video sequences, exploiting the general principle that the main function of the human visual system is the extraction of structural information from the viewing field. The SSIM was selected to be used for the experimental needs of this paper, due to its satisfactory evaluation performance in the relative performance evaluation studies [9], [10]. Thus, considering that f and f' depicts the frames of the uncompressed and compressed signal respectively, then the SSIM is defined as:

$$SSIM(f,f') = \frac{(2\mu_f \mu_{f'} + C_1)(2\sigma_{ff'} + C_2)}{(\mu_f^2 + \mu_{f'}^2 + C_1)(\sigma_f^2 + \sigma_{f'}^2 + C_2)} \qquad (2)$$

where μ_f, $\mu_{f'}$ are the mean of f and f', σ_f, $\sigma_{f'}$, $\sigma_{ff'}$ are the variances of f, f' and the covariance of f and f', respectively. The constants C_1 and C_2 are defined as:

$$C_1 = (K_1 L)^2 \quad C_2 = (K_2 L)^2 \qquad (3)$$

where L is the dynamic pixel range and $K_1 = 0.01$ and $K_2 = 0.03$, respectively.

IV. VIDEO CODING OF HEVC AND AVC SIGNALS

For the evaluation process two reference video clips (Bubbles, Horse_Race) [11] and two non-reference videos (Apocalypto_Trailer and Batman_Dark_Night_Trailer) were used, which represent various spatiotemporal levels. A representative frames of each signal are depicted on Figure 2.

Fig. 2. Representative frames taken from the test signals

The video clips were encoded from their original uncompressed YUV format to ISO AVC Main Profile (MP) and to the following profiles of HEVC, namely: (1) Random Access Profile (RAP), (2) Random Access Low Complexity Profile (RALCP), (3) Low Delay Profile (LDP) and (4) Low Delay P Profile (LDPP). Across all the encoding process, the reference software was used for both the AVC and HEVC coding. For HEVC the Test Model (HM) Reference Software v5.1 [12] used and similarly for AVC the JM Reference one.

For reference reasons, all the test signals, which were used in this paper, are available for downloading at [13]. Additionally are provided for future reference the AVC and HEVC signals, which were used for this paper. The data of the test signals are provided on Table I.

In order to achieve an accurate comparison between the various profiles, it is necessary all the rest configurations of the codec to be either identical or if this is not possible then with very similar parameters or values. For this reason, the GOP structure for all the encoding profiles and between the two encoding methods consisted of I, P and B frames, except in cases of LDP and LDPP where I/B and I/P patterns were used respectively, ensuring the benchmarking of both Intra- and Inter-coding efficiency between AVC and HEVC profiles.

TABLE I. TEST SIGNALS DATA

Test Signal	Frames	Resolution
Apocalypto Trailer	990	352x288
Batman Trailer	913	352x288
Bubbles	501	416x240
Horses	300	416x240

The Quantization Parameter (QP) has a great impact on visual quality and compression ratio, as it regulates how much spatial detail is maintained. For this reason, a variety of QP values {12, 22, 32, 42, 51} were used for all the experimental part, creating a more complete set of measurements.

V. BENCHMARKING OF HEVC AND AVC SIGNALS

A. Performance Comparison of HEVC vs. AVC

In this sub-section, the coding performance of the HEVC algorithm is examined in comparison to the AVC. During the evaluation test the same encoding parameters were selected for both codecs (i.e., the QP for both I and P or B frames were set sequentially to the values {12, 22, 32, 42, 51}.

TABLE II. AVERAGE PSNR OF AVC (MP) VS. HEVC

QP	MP	LDP	LDPP	RAP	RALCP
12	52.08022	50.07781	50.14053	48.60677	48.04049
22	43.82217	42.21041	42.13275	41.87272	41.37249
32	36.71928	34.82381	34.79465	35.04391	34.73854
42	29.62316	29.04746	29.03886	29.51388	29.30517
51	11.73947	24.85184	24.86316	25.35809	25.20807

The first comparison between HEVC and AVC is related to the average PSNR vs. QP value. Table II shows the average PSNR values and the corresponding QPs for all the test signals encoded at the AVC Main Profile (MP) and the four HEVC profiles (LDP, LDPP, RAP, RALCP). The graphical representation of the PSNR values is depicted on Figure 3.

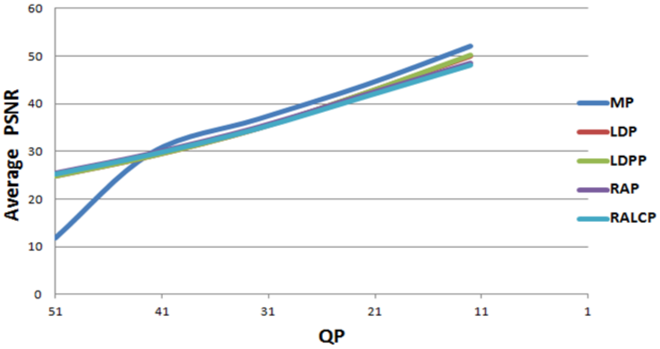
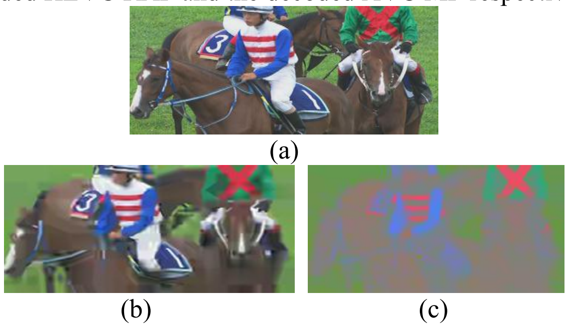


Fig. 3. Average PSNR vs QP of HEVC/AVC signals.

Based on the graphical representation, it can be observed that, the HEVC profiles have slightly degraded performance efficiency (i.e., 1-2dB), as compared to the Main Profile of AVC for QP values between 12 and 42. Furthermore AVC achieves slightly better performance on low QPs, however as the QP value gradually increases, the AVC's encoding efficiency is reduced, reaching to an abrupt downfall when the QP equals to 51. On the other hand HEVC outperforms AVC as it scores PSNR at acceptable values, for QP higher than 42, while retaining its linearity despite the QP reduction. Additionally, the AVC's decoded video for QP equal to 51, proves not viewable, while HEVC achieves acceptable video quality and smooth playback. This is depicted in Fig. 4, which shows a representative frame of the original signal, the decoded HEVC RAP and the decoded AVC MP respectively.


(a)

(b) (c)

Fig. 4. Decoded frame of Race Horses signal for HEVC RAP and AVC MP (QP=51) (a) Initial Video, (b) HEVC RAP, (c) AVC MP

B. Compression Comparison of HEVC vs. AVC

This sub-section presents the experimental results on the compression efficiency between HEVC and AVC. During the evaluation tests the same set of QP values was used as previously. The average value of bit rate (i.e., bps) at each encoding profile for the whole duration is shown at Table III.

TABLE III. AVERAGE BITRATE VS. QP FOR AVC (MP) AND HEVC

QP	MP	LDP	LDPP	RAP	RALCP
12	10716.83	5080.18	5402.37	4176.73	4011.80
22	2054.20	1249.34	1316.25	1105.97	1083.83
32	582.20	284.73	286.45	274.38	277.29
42	114.24	78.63	77.82	78.32	77.70
51	6.80	34.10	33.65	34.58	33.25

The graphical representation of Table III is presented in Fig. 5, while in Fig. 6 a zoom in is shown especially for the area of interest (i.e. the low bit rate cases). In Fig. 5, the apparent gap can be noticed separating the AVC average bit rate curve from the HEVC ones, in a distant view including values up to 12 kbps. Moreover, a differentiation in the performance among the HEVC profiles is also noticeable.

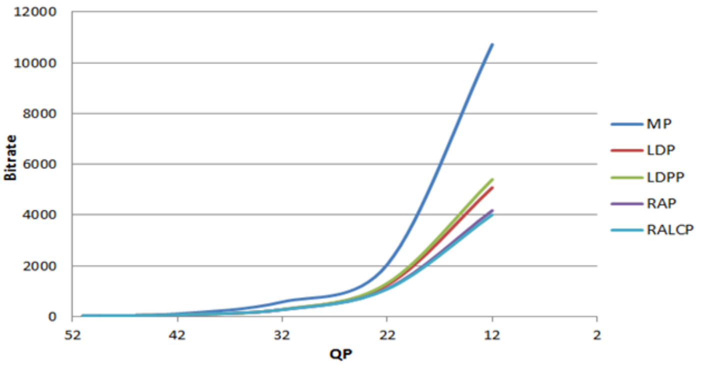


Fig. 5. Average Bitrate vs. QP for AVC (MP) and HEVC

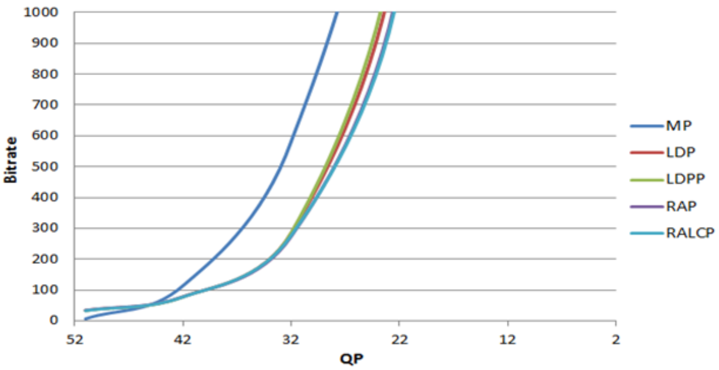


Fig. 6. Zoom in the area of 0-1 kbps of the Average Bit rate curves vs. QP for AVC (MP) and HEVC (LDP, LDPP, RAP, RALCP).

Additionally in Fig. 6, a more detailed and zoomed in view of the same evaluation tests can be observed, portraying the compression efficiency of the two codecs in low bit rate cases, even better. However, for QP values between 42 and 51, AVC surpasses HEVC in compression ratio, achieving significantly lower bit rate values (i.e., higher compression ratio). But as, it is thoroughly depicted in the previous section (i.e. Fig. 2) the performance efficiency of AVC, in QP value equal to 51, is extremely low, making the decoded content not perceptually accepted. Therefore, although the compression is higher, the resulted signal is perceptually unacceptable, making the specific QP area not operational for AVC case.

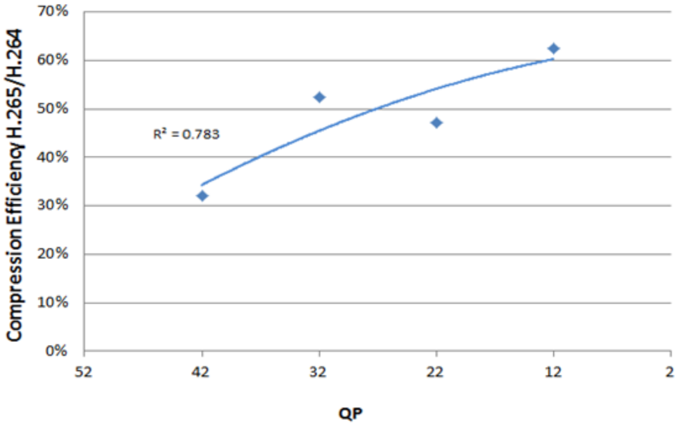


Fig. 7. Compression Efficiency (%) vs. QP for HEVC RALCP-AVC MP.

A more comprehensive representation of the compression efficiency (%) of HEVC in comparison to AVC is depicted in Fig. 7. The curve of this figure shows that for QP values from 12 to 42, the compression efficiency improvement of HEVC ranges from 32% to 62%. For values of QP between 42 and 51 the curve is intentionally not considered, since for these values the AVC-coded signal is significantly degraded.

C. Video Quality Comparison of HEVC vs. AVC

In this sub-section, the video quality of the HEVC algorithm is compared to the AVC case for the video signals under test. In this test, as well as previously, the same encoding parameters were retained for the whole encoding procedure (i.e. the same QP for both I and P or B frames were used).

TABLE IV. AVERAGE SSIM VS. QP FOR AVC (MP) AND HEVC

QP	MP	LDP	LDPP	RAP	RALCP
12	0.997645	0.996220	0.996283	0.994878	0.994350
22	0.986573	0.982075	0.981653	0.982208	0.980438
32	0.941283	0.925773	0.925538	0.934403	0.930885
42	0.811995	0.798598	0.798740	0.818493	0.815078
51	0.421673	0.665193	0.666523	0.687155	0.685240

Initially the average SSIM was calculated for all the test signals under the AVC (MP) and the HEVC profiles (LDP, LDPP, RAP, RALCP). Table IV contains the experimental results and Fig. 8 depicts their graphical representation.

It can be observed that, the HEVC and AVC profiles have similar video quality performance for all the QP values from 42 to 12. However, for higher QP values (i.e., between 42 and 51), the video quality of the AVC profile collapses to unaccepted levels, while it is not noticed any significant variation at the video quality of the HEVC profiles, which retains its linearity despite the QP reduction.

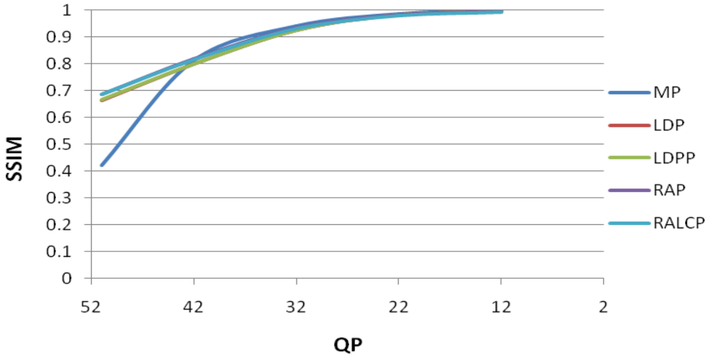


Fig. 8. Average SSIM vs. QP for HEVC and AVC profiles

Therefore, in conjunction with the results of the previous sub-sections, it deduced that that the HEVC can achieve the same video quality of AVC coding for values of QP between 42-12, while for the same range of QP values, the improvement of the compression and encoding performance of HEVC over AVC is 32% to 62%.

VI. CONCLUSIONS

This paper presented a quantitative and qualitative comparison between the HEVC and AVC performance with scope to show if the objectives of the emerging standard have been achieved. Based on experimental data objectively evaluated and estimated, it is shown that the emerging HEVC can retain the same video quality level as AVC, while HEVC also achieves 32% to 62% improvement in the compression and coding efficiency, depending on the selected QP value. The achieved improvement in the encoding efficiency is mainly resulted from the use of extended block sizes in the coding process and also from the recursive quadtree approach, which can adapt the accuracy and performance of the coding algorithm on the complexity of the content. Also it is shown that the encoding enhancement of HEVC over AVC is not constant but varies on the value of the QP and the type of the video content. Finally, for very high values of QP, the HEVC continues to operate smoothly (maintaining the video quality at accepted levels), while the AVC for the same QP values could not encode efficiently. This paper concludes that the HEVC encoder has managed to achieve the objectives and goals of the emerging standard in terms of performance, making HEVC/H.265 an appropriate standard to further support the future content distribution systems.

REFERENCES

[1] ISO/IEC International Standard 14496-10, Information Technology – Coding of Audio-Visual Objects – Part 10: Advanced Video Coding, third ed., Dec. 2005, corrected version, March 2006.

[2] G. J. Sullivan and T. Wiegand, “Video compression – from concepts to the H.264/AVC standard,” Proceedings of the IEEE, vol. 93, no. 1, pp. 18-31, Jan. 2005.

[3] F. Pereira and T. Alpert, “MPEG-4 Video Subjective Test Procedures and Results”, IEEE Transactions on Circuits and Systems for Video Technology. Vol. 7(1), pp. 32-51, 1997.

[4] Kemal Ugur, Kenneth Andersson, Arild Fuldseth, Gisle Bjøntegaard, Lars Petter Endresen, Jani Lainema, et al., "High Performance, Low Complexity Video Coding and the Emerging HEVC Standard", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 20, No.12, December 2010

[5] Ken McCann, Woo-Jin Han, Il-Koo Kim, Video coding technology proposal by Samsung and BBC, JCTVC-A124, JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, October, 2010

[6] T. Wiegand, H. Schwarz, BoG report: Residual Quadtree Structure, JCTVC-C319, JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, October, 2010

[7] B. Bross, H. Kirchhoffer, H. Schwarz, T. Wiegand, Fast intra encoding for fixed maximum depth of transform quadtree, JCTVC-C311, JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, October, 2010

[8] [8] J. L. Martínez, P. Cuenca, F. Delicado, F. Quiles, “Objective Video Quality Metrics: A Performance Analysis”, 14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006

[9] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Proc. 13 (4) (2004) 600-612.

[10] Z. Wang, L. Lu, A. C. Bovik, Video quality assessment based on structural distortion measurement, Signal Processing: Image Comm. 19 (2) (2004) 121-132.

[11] Link for the reference test video signals ftp://ftp.tnt.uni-hannover.de/testsequences

[12] HEVC reference toolkit available for download version directory https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/

[13] YUV uncompressed and decoded signals available for downloading on http://aias.iit.demokritos.gr/~koumaras/HEVC/yuv_signals.html