HEVC:A Review, Trends and Challenges

Antonio J. Díaz-Honrubia[1], José Luis Martínez[1], Pedro Cuenca[1]

Abstract— **HEVC is a draft of a codec which is currently been developed jointly by the ISO/IEC and the ITU-T and it is expected to be the natural successor of H.264/AVC. When completed, it will be suitable for resolutions up to Ultra High Definition video coding, which corresponds with 7680x4320 pixels. Moreover, it can achieve better quality results at the same bitrate against its predecessor H.264/AVC, which makes very probable that HEVC will be the most extended video coding standard in a near future. This paper aims to evaluate the HEVC coding efficiency and, thus, establish the hottest research lines focusing on speeding-up the encoding time and the H.264/HEVC transcoding.**

Keywords— **HEVC, H.264/AVC, performance, challenges**

I. Introduction

NOWADAYS, *H.264* or *Advance Video Coding* (AVC) [1] is the most extended video compression standard for High Definition (HD) video coding. Thus, the most important applications are the HD Television (HDTV) and the Bluray video storage. As a consequence of this fact, there are a lot of contents according to this standard.

Nevertheless, the ISO/IEC Moving Experts Group (MPEG) and the ITU-T Video Coding Experts Group (VCEG) joint their efforts in January 2010 as the Joint Collaborative Group on Video Coding (JCT-VC) to develop a new standard, which was pretended to be the natural evolution of H.264/AVC. Then, they issued a formal Call for Proposals (CfP) and in April of that same year they evaluated a set of 27 proposals. After a performance and visual quality evaluation [2] (which is said by the authors to be *the largest video subjective quality testing effort ever conducted*) some proposals showed that could achieve about a30% of the bit rate reduction at the same visual quality than H.264/AVC.

However, the price to pay for this bit rate reduction is a more complex video coding algorithm with computationally more expensive tools. In fact, the proposals multiplied the computational cost of H.264/AVC in a factor from 2 to 10. At present, this new standard is still under development and it is pretended to be ready in early 2013. Its name will be *High Efficiency Video Coding* (HEVC), but it is also known as *H.265* and *MPEG-H Part 2*.The Reference Software for HEVC codec, *HM* (*HEVC test Model*), contains those techniques which are more promising to be included in the final standard and which are included in the current working draft of the standard [3].

In this paper a HEVC performance evaluation is going to be carried out to be able to appreciate the computational cost of the coder. Moreover, the main lines of research are shown, such as H.264/AVC to HEVC content adaptation or HEVC encoding time reduction.

The remainder of this paper is organized as follows: Section II includes a technical background of H.264/AVC and HEVC coding tools, Section III shows the trends and related work which is being developed about the topic, Section IV provides a comparison of the two video coding standards and, finally, challenges and conclusions are shown in Section V.

II. Technical Background

In order to know the main differences among H.264/AVC and HEVC, a technical background with the main coding tools of each standard is presented here. As it can be seen in this section, HEVC is based on the same hybrid spatial-temporal prediction system as its predecessor H.264/AVC (which was also based on its precursor codecs, such as MPEG-2 [4]).

A. H.264/AVC

This standard introduced several changes regarding the previous video coding standards which resulted in a significant bitrate reduction at the same visual quality [5]. There were two main novelties in the standard: the Inter and the Intra Prediction. The first one introduced a more flexible partition of a Macroblock (MB) to carry out the Motion Estimation (ME), so each 16x16 pixels MB could be partitioned into 8x16, 16x8 or 8x8 pixels sub-MBs and, if the 8x8 was selected, each sub-MB could be partitioned into 4x8, 8x4 or 4x4 pixels sub-MBs. On the other hand, Intra Prediction was not exist in previous standard and it is a completely novelty of H.264/AVC. It consisted in extrapolating already decoded neighbouring pixels to predict the present block. Each block could have a size of 16x16, 8x8 or 4x4 pixels and it could be predicted using a whole of 9 predictors (only 4 of these modes are available for the 16x16 pixels block). These prediction modes are shown in Fig. 1 (except mode 2, which consists in extrapolate the value by the average of the predictors).

Other changes introduced by H.264/AVC are the use of an integer transform, which can be applied in blocks of 8x8 or 4x4 pixels, the possibility to use *Context-Adaptive Binary Arithmetic Coding* (CABAC) as

[1] Instituto de Investigación en Informática de Albacete (I3A). Universidad de Castilla-La Mancha (UCLM). Campus Universitario de Albacete. 02071. Albacete (SPAIN).

entropy coder and the introduction of a deblocking filter whose objective is to reduce the blocking effect caused by the different modes selected for each MB coding.

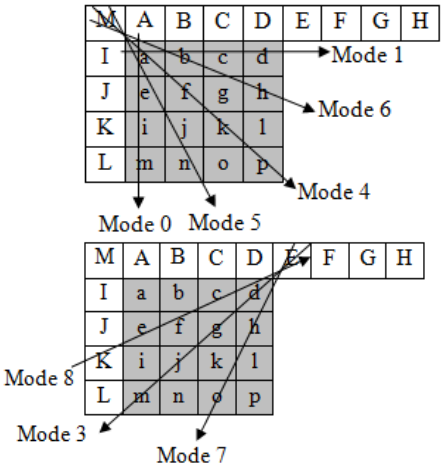


Fig. 1. Intra Prediction directional modes for a 4x4 block in H.264/AVC

B. HEVC

This paradigm also introduces some new coding tools, as well as it improves other already used tools [6]. Maybe, the most important change affects to the picture partitioning. HEVC dispense with the terms of MB and block for the ME and the transform respectively and introduces three new concepts: Coding Unit (CU), Prediction Unit (PU) and Transform Unit (TU). This structure leads to a more flexible coding to suit the particularities of the image. Each picture is partitioned into squared regions of variable size called CUs which replace the MB structure of previous standards. Each CU may contain one or several PUs and TUs and its size is limited from 8x8 to 64x64 pixels. To fix the size of each CU, first of all, a picture is divided into 64x64 pixels areas, each of which is called Largest CU (LCU), and then, each LCU can be partitioned into 4 smaller sub-areas of a quarter of the original area. This partitioning can be done with each sub-area recursively until it has a size of 8x8 pixels. Thus, a quadtree structure is used, as it can be seen in Fig. 2.

A PU is the elementary unit for prediction (as sub-MB in H.264/AVC) and they are defined at quadtree leaves, i.e., the largest PU size is equal to the CU size. However, a PU could be smaller than the CU depending on the prediction type selected for it. If the PU is skipped, the only possibility is to have the same size than the CU. If it is Intra Predicted it can have the same size or to be partitioned into 4 square PUs of the same size. Finally, if it is Inter Predicted, there are 8 different partitions: preserve the size, split it in 2 rectangular PUs of the same size horizontally or vertically, split it in 4 square PUs and split it in two asymmetric rectangular PUs, where a smaller PU covers 1/4 of the CU and a bigger PU cover 3/4 of the CU. This last mode are, really, 4 four modes, depending on whether the splitting is done horizontally or vertically and whether the smaller PU is on the left/bottom or on the right/top. This asymmetric partitioning is a novelty of HEVC and it aims to detect the borders in the picture to carry out a better ME. Finally, a TU is the basic unit for transform and quantization and, as s PU, it is defined at the quadtree leaves. A TU can have a size from 4x4 to 32x32 pixels and can contain one or more PUs. If in the current CU there are not asymmetric partitions of the PUs, the TU can have the same size as the CU (unless it is a 64x64 CU) or be split into 4 equal TUs. If the CU contains asymmetric PUs, then the TU can have the same size as the CU or be split into 16 equal TUs.

Other main difference is about Intra Prediction, where, now, a whole of 35 modes are available to select for prediction: 33 angular directions, DC mode and Planar mode [7].

HEVC checks each almost all these predictions modes (Inter and Intra) to decide whether it splits a CU/PU/TU or not and it chooses the case which produces the best Rate-Distortion. This wide range of possibilities makes HEVC to be much more computationally expensive than its predecessor H.264/AVC.

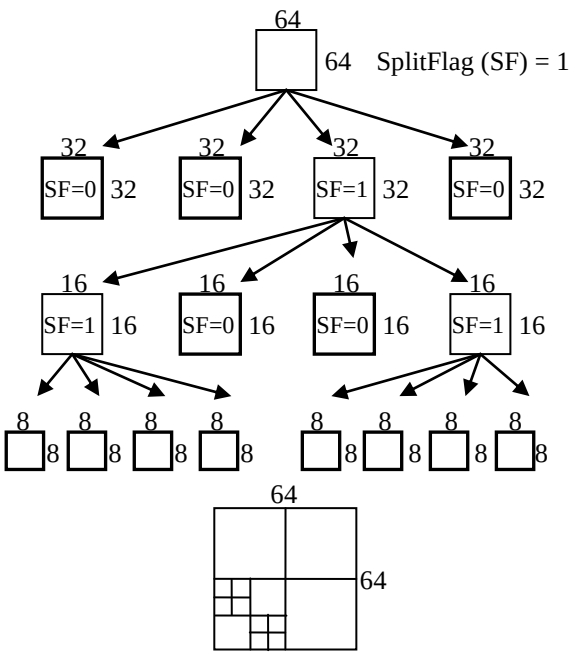


Fig. 2. CU splitting with its representation in a quadtree structure

III. Trends and Related Work

In the last years, the introduction of small devices which captures and reproduces video, the increase in TV sizes and the demand of better quality by society have induced a change in video coding technology. Therefore, nowadays, one of the trends is to use bigger resolutions for video coding. In these bigger pictures, more pixels represent the same real area than in smaller pictures, making prediction areas can be expanded. Moreover, small devices still make necessary the use of small prediction areas, so the bigger blocks should be able to be split. This was taken into account in most answers to the CfP released in 2010. Although one of the best performance proposals used an 8x8 basic block for Inter Prediction, the approach was not finally included because of the above.

HEVC coding, as shown in Section IV, takes more time than previous standards and this produces more power consumption. Because of this, optimization is required. A lot of work has been carried out to reduce ME and picture partitioning complexity in previous standards. For example, in [8] a dynamic ME method is proposed for H.264/AVC, which uses full fast search or several search techniques. Furthermore, [9] proposes an algorithm, also for H.264/AVC, to adjust the number of coding modes according to time requirements of any application. However, no related work has been found for HEVC. Besides, some efforts are being carried out to develop a H.264/AVC to HEVC transcoder, due to the wide use of H.264/AVC nowadays [10]. However it is a very open field which should be deeper researched, as the CU, TU and PU to MB conversion, the selection of an appropriate Intra mode or the use of Graphic Processing Units (GPUs) to accelerate the transcoding.

Finally, other interesting trend is the use of Multiview Video (MV) with HEVC. There is a lot of work made about MV for H.264/AVC which could be easily extrapolated to HEVC. Some proposals only include some views with Inter references between them [11], while others include compressed depth maps with each view, which makes possible to create more virtual views [12]. In [13] an approach to integrate MV into HEVC can be seen and it achieves a reduction of a 37,2% in the bitrate compared to simulcast HEVC views coded independently.

IV. Performance Evaluation

As an effort to carry out a good evaluation of the standard, the JCT-VC developed a document with some reference sequences and the codec configuration which should be used with each one [14]. The sequences are divided into 6 groups. The A group corresponds with sequences with a resolution of 2560x1600 pixels, the B group contains sequences of 1080p, the C group consists of sequences with a resolution of 832x480 pixels, the D group sequences have a resolution of 416x240 pixels and the E group consists of 720p sequences. The last group, F, is a combination of multiple resolutions.

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 this document. A sequence from each class has been selected for these simulations, all of them with an 8 bit depth representation:

- Class A: *Traffic* (2560x1600 pixels). 150 frames at 30 frames per second (fps).
- Class B: Cactus (1920x1080 pixels). 500 frames at 50 fps.
- Class C: *BasketballDrill* (832x480 pixels). 500 frames at 50 fps.
- Class D: *BlowingBubbles* (416x240 pixels). 500 frames at 50 fps.
- Class E: *FourPeople* (1280x720 pixels). 600 frames at 60 fps.
- Class F: *ChinaSpeed* (1024x768 pixels). 500 frames at 30 fps.

All sequences have been simulated with two profiles, three configurations for each profile and QP values of 22, 27, 32 and 37 for each profile/configurations pair. Profiles are Main and High Efficiency and configurations are Random-Access (class E sequences does not use this first configuration), Low-Delay with P frames and Low-Delay with B frames (class A sequences does not use these two last configurations). All these profiles, configurations and QP values are defined in [14].

Simulations have been coded using H.264/AVC standard and HEVC working draft 6. The coding tools are the reference encoders made by the JCT-VC: JM 18.3 for H.264/AVC and HM 6.1 for HEVC. Both include configuration files which are prepared for coding with the profiles and configurations previously mentioned, hence, the only parameter which has been modified is the QP. In the case of H.264/AVC, there is not possibility of configure it with HEVC-like Main profile and Low-Delay with B frames configuration, so simulations have only been launched with high efficiency profile and Random-Access and Low-Delay-P configurations with JM. To carry out the comparisons, the results are calculated as indicated in [15], in terms of bit rate saving and PSNR increment for the luminance component. It is known that PSNR is not a perfect metric for video quality, as it cannot measure the subjective quality perceived by the viewers, but it is easier to measure and good enough to satisfy the aim of this paper. Average time increase percentage has been also included.

A. H.264/AVC and HEVC comparison

It is important to know the computational cost of HEVC compared to H.264/AVC to know whether the new codec can be implemented in the existing coding devices. Thus, a comparison between JM 18.3 and HM 6.1 has been performed with those configurations which are available in both coders: High Efficiency profile with Random-Access and Low-Delay-P configurations.

TABLE I. HEVC GAIN AGAINST H.264/AVC

HIGH EFFICIENCY PROFILE AND RANDOM-ACCESS CONFIGURATION

Sequence class	ΔPSNR (dB)	Bitrate saving (%)	ΔCoding Time (%)
A	4,06	72,07	74,60
B	1,12	12,59	20,24
C	3,44	69,60	30,59
D	2,50	64,78	12,45
F	1,61	75,16	45,33
Average	***2,28***	***58,84***	***36,64***

As it can be seen in Table I, in the Random-Access case, HEVC gets an average gain of 2,28 dB, while reducing the bitrate in a 58,84%. These achievements are obtained by using more complex coding tools, as the quadtree structure for prediction or a bigger predictor's set for Intra Prediction. However, the use of these tools implies an increment of 36,64% in the coding time. Results for the Low-Delay-P case, which are shown in

Table II, are equals to the previous ones, except the bitrate saving, which is lower in this case. This result may suggest that HEVC improves Intra Prediction more than Inter Prediction, since it achieves more bitrate reduction when using I frames (in the Low-Delay configurations only the first frame is an I frame, while in the Random-Access configuration an I frame appears every 32 frames.

TABLE II. HEVC GAIN AGAINST H.264/AVC HIGH EFFICIENCY PROFILE AND LOW-DELAY-P CONFIGURATION

Sequence class	ΔPSNR (dB)	Bitrate saving (%)	ΔCoding Time (%)
B	1,67	35,99	54,16
C	2,55	40,68	32,59
D	2,39	30,63	35,98
E	2,33	41,28	74,56
F	2,46	31,99	41,49
Average	***2,28***	***36,11***	***47,76***

B. HEVC profiles comparison

After the comparison with H.264/AVC, in this sub-section High Efficiency and Main HEVC profiles are compared. The differences between these profiles are that the High Efficiency one enables adaptive loop filter, intra chrominance prediction based on reconstructed luminance, non-square transforms, asymmetric motion partitions and the use of 10 bits for internal representation of a pixel.

As it can be seen in Table III, where the Random-Access configuration has been used to compare the profiles, a gain of 0,19 dB is obtained, while reducing the bitrate almost a 5%. However, the more complex tools used in the High Efficiency Profile, implies an increasing of the 17% in the coding time. Results in Table IV and Table V, where the Low-Delay-P and the Low-Delay-B configurations have been used, are pretty similar to the Table III ones, what points out that the use of different configurations does not influence in the gain of using High Efficiency Profile instead of Main Profile. Due to the coding time increment, it does not seem logical to use High Profile in real time scenarios, but it could achieve better results at films coding, where no requirement time are taken into account.

RD curves for each sequence A to F can be seen in Fig. 3. As can be appreciated, all HEVC simulations get better results than H.264/AVC ones. Among HEVC profiles and configurations can be observed that High Efficiency Profile with Random-Access configuration is the best-performance case and Main Profile with Low-Delay-P configuration is the worst one as resulting from the use of lower complexity tools.

Furthermore, in this figure it can be seen that Main Profile with Low-Delay-B configuration has similar quality than High Efficiency Profile with Low-Delay-P configuration and the same occurs with Main Profile with Random-Access configuration and High Efficiency Profile with Low-Delay-P.

TABLE III. HEVC MAIN PROFILE AGAINST HIGH EFFICIENCY PROFILE RANDOM-ACCESS CONFIGURATION

Sequence class	ΔPSNR (dB)	Bitrate saving (%)	ΔCoding Time (%)
A	0,26	7,11	13,18
B	0,18	6,64	14,45
C	0,20	4,85	18,72
D	0,15	3,34	18,67
F	0,15	3,00	20,77
Average	***0,19***	***4,99***	***17,16***

TABLE IV. HEVC MAIN PROFILE AGAINST HIGH EFFICIENCY PROFILE LOW-DELAY-P CONFIGURATION

Sequence class	ΔPSNR (dB)	Bitrate saving (%)	ΔCoding Time (%)
B	0,20	6,58	14,79
C	0,35	6,67	17,57
D	0,08	1,78	20,64
E	0,13	2,47	17,47
F	0,34	9,07	8,71
Average	***0,22***	***5,31***	***15,84***

TABLE V. HEVC MAIN PROFILE AGAINST HIGH EFFICIENCY PROFILE LOW-DELAY-B CONFIGURATION

Sequence class	ΔPSNR (dB)	Bitrate saving (%)	ΔCoding Time (%)
B	0,15	5,08	13,01
C	0,25	5,91	17,12
D	0,14	2,97	21,06
E	0,33	8,32	7,69
F	0,16	3,02	16,70
Average	***0,21***	***5,06***	***15,12***

V. CONCLUSIONS AND CHALLENGES

The main conclusion of this paper is that HEVC obtains the same quantitative quality than H.264/AVC at approximately the 50% of bitrate saving. However, as more complex tools are used, coding time is incremented in a 40%. This gain is a consequence of changes as the use of larger Prediction Units than Macroblocks heretofore used, the improvement of Intra Prediction with more prediction modes and directions or the use of asymmetric partitions in Inter Prediction.

However, this gain occurs at expense of an increase of the coding time, especially at more complex profiles and configurations, as previously said. This makes HEVC non-viable for real-time applications, such a multimedia server, where there are lots of videos coded with H.264/AVC standard and they could be transmitted to a HEVC decoder, making real-time transcoding necessary. This could easily happen because a lot of content coded with H.264/AVC standard already exists. Regarding this, a lot of work remains to be done, since some tools are pretty different in both standards, such as the picture splitting or the substantial increase of search options. So, one of the main challenges nowadays is to design a H.264/AVC to HEVC transcoder. In this point, firstly a design of the architecture for the transcoder is needed, studying the H.264/AVC decoder and the HEVC

encoder and noting the possible parallelisms and differences among the standards.

After this preliminary study, an acceleration of the coding part of the transcoder should be carried out, since H.264/AVC decoder works in real time. Traditionally, the most expensive part in a coder is the ME, so this module of the HEVC coder is a good objective to reduce the complexity. Moreover, HEVC increases the complexity because the CU, TU and PU partitioning in the quadtree structure is very expensive, so an acceleration of this part, of which only the TU partitioning belongs to the ME module, could be interesting to be done.

To accelerate the partitioning, information of the H.264/AVC decoding could be used, extrapolating the unit sizes from the blocks and MBs sizes used by the original standard, since the partitioning in H.264/AVC can reflect the characteristics of the original image. Furthermore, it could also help to enable or disable some partitions depending on the picture characteristics. However, HEVC uses larger block sizes and it could be a problem. This problem could be solved using an additional way to split the picture, as the statistical information which is contained in the residual. For instance, the more bits are needed to code the residue, the more movement there is in the original picture. Another idea to solve the larger block size is to use a clustering algorithm to join some MBs in the residue to produce larger CUs, PUs and TUs.

Regarding the ME acceleration, starting with the motion vectors of the H.264/AVC coded sequence, HEVC coder can be accelerated using that motions vectors as an approximation of the new ones. Thus, the new coder would only have to search in a reduced area using as the coordinate origin the position where the previous vector points.

Finally, one of the most open research fields nowadays is the multiview video coding, since it was the last addition to the H.264/AVC and it is expected to add a multiview extension for HEVC. It is not decided whether this extension will use only video layers or video and depth layers. Until that moment, little work can be advanced, but when this is sure, acceleration in this field will be required. Besides, H.264/AVC was the first standard including a multiview extension, so there are already contents coded with that standard and transcoding will be needed too.

ACKNOWLEDGMENTS

This work was supported by the Spanish MEC and MICINN, as well as European Commission FEDER funds, under Grants CSD2006-00046 and TIN2009-14475-C04. It was also partly supported by The Council of Science and Technology of Castilla-La Mancha under Grants PEII09-0037-2328. Finally, it has also been supported by the “Red Temática en Codificación y Transmisión de Contenidos Multimedia (RTCTCM)” of the Spanish “Ministerio de Ciencia e Innovación (TEC2010-11776-E)”.

REFERENCES

[1] ISO/IEC, *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding.* ISO/IEC International Standard, 2005.

[2] T. Wiegand, J.R. Ohm, G. Sullivan, W. J. Han, R. Joshi, T. K. Tan and K. Ugur, *Special Section on the Joint for Proposals on High Efficiency Video Coding (HEVC) Standarization*, IEEE Transactions on Circuits and Systems for Video Technology, December 2010.

[3] Joint Collaborative Team on Video Coding of ITU-T and ISO/IEC, *High efficiency video coding (HEVC) text specification draft 6*, 8th Meeting of the JCT-VC in San José (USA), February 2012.

[4] ISO/IEC, *MPEG-2: Generic coding of moving pictures and associated audio information*, Document 13818-1, 2000.

[5] T. Wiegand, G.J. Sullivan, G. Bjøntegaard, A. Luthra, *Overview of the H.264/AVC Video Coding Standard*, IEEE Transactions on Circuits and Systems for Video Technology 13 (7) (2003) 560–576.

[6] O. Le Meur, *Video Compression Beyond H.264, HEVC*, Université de Rennes, November 2011.

[7] J. Lainema, K. Ugur and O. Bici, *Planar Intra Coding for Improved Subjective Video Quality*, 4th Meeting of the JCT-VC in Daegu (South Korea), January 2011.

[8] C. Man-Yau and S. Wan-Chi, *Computationally-scalable motion estimation algorithm for H.264/AVC video coding*, IEEE Transactions on Consumer Electronics, vol. 56, pp. 895-903, 2010.

[9] R. Jianfeng, N. Kehtarnavaz, and M. Budagavi, *Computationally efficient mode selection in H.264/AVC video coding*, IEEE Transactions on Consumer Electronics, vol. 54, pp. 877-886, 2008.

[10] Eduardo Peixoto, Ebroul Izquierdo, “A Complexity-Scalable Transcoder form H.264/AVC to the New HEVC Codec”, to appear in International Conference on Image Processing (ICIP), Orlando, FL, USA, September 2012

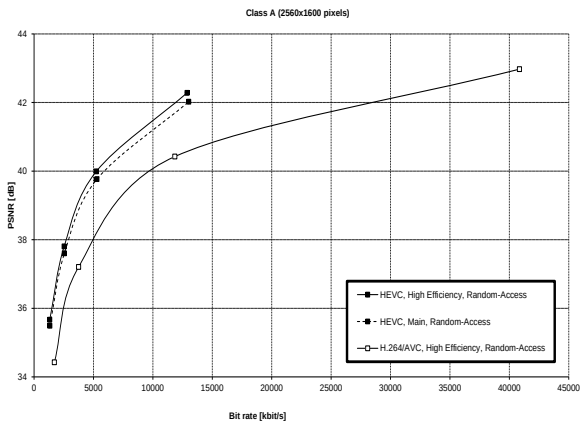
[11] P. Merkle, A. Smolic, K. Müller and T. Wiegand, *Efficient Prediction Structures for Multi-view Video Coding*, 2007.

[12] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Mueller, P.H.N. de With and T. Wiegand, *The Effects of Multiview Depth Video Compression on Multiview Rendering*, SignalProcessing: Image Communication, 2009.

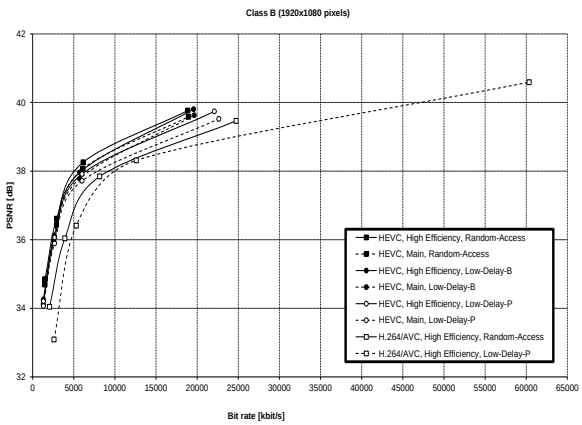
[13] G. Van Wallendael, S. Van Leuven, J. De Cock, F. Bruls and R. Van de Walle, *3D Video Compression Based on High Efficiency Video Coding*, 2012.

[14] F. Bossen. *Common test conditions and software reference configurations*, 9th Meeting of the JCT-VC in Geneva (Switzerland), April 2012.

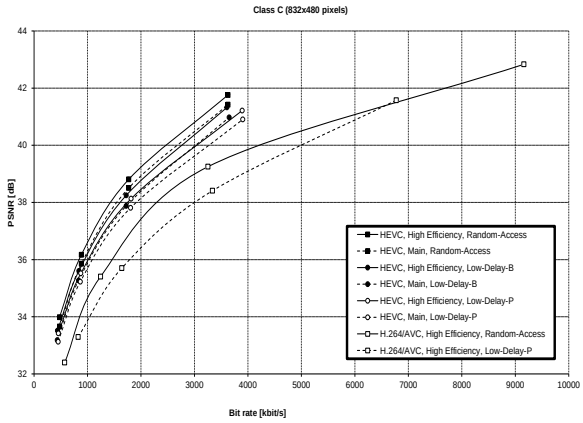
[15] G. Bjøntegaard, *Calculation of Average PSNR Differences between RD-Curves*, 13th Meeting of the VCEG-M33 in Austin (Texas, USA), April 2001.



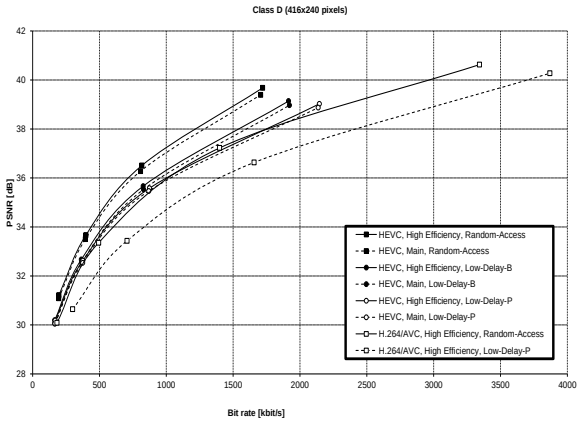
a. Traffic sequence



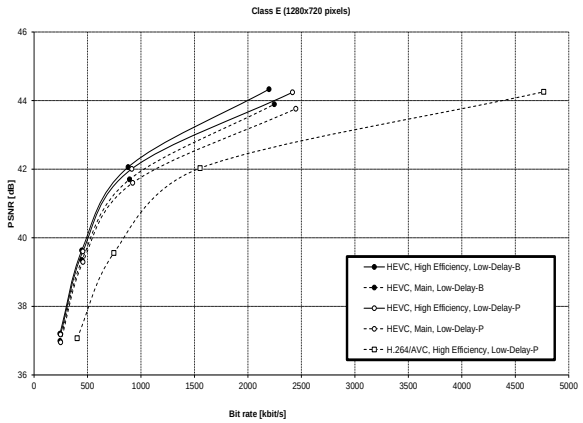
b. Cactus sequence



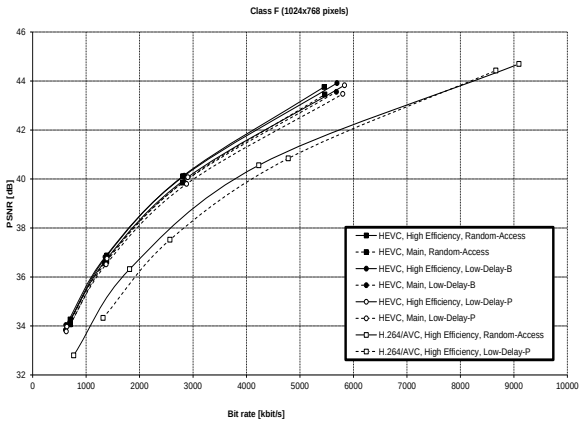
c. BasketballDrill sequence



d. BlowingBubbles sequence



e. FourPeople sequence



f. ChinaSpeed sequence

Fig. 3. RD graphics for each simulation launched during performance evaluation