Comparison between HEVC and Thor based on objective and subjective assessments

Comparison Between HEVC and Thor Based on Objective and Subjective Assessments

Lorenzo Re Fiorentin[1], Maurizio Masera[1], Sabino Mantovano[2], Paola Sunna[2], Guido Masera[1], Maurizio Martina[1]
[1]Politecnico di Torino, 10129 Turin, Italy
[2]CRIT RAI, 10129 Turin, Italy
{*name.surname*}@*polito.it,*{*name.surname*}@*rai.it*

***Abstract*—The increasing popularity of HD TV, the emergence of beyond-HD formats and higher quality and resolution mobile TV services is posing a severe challenge on today's broadcast networks. Next generation networks (e.g, DVB-T2, DVB-S2, etc.) and higher coding efficiency can provide a solution to this problem. High Efficiency Video Coding (HEVC) is the latest standard jointly released by the ITU-T Video Coding Experts Group and ISO/IEC Moving Picture Experts Group (MPEG), but other options, royalty free, such as Thor, have recently emerged and are gaining popularity. In this paper a rate-distortion analysis of HEVC and Thor, is presented. The comparison is performed using both objective and subjective results and the results indicate an overall better performance of HEVC.**

***Keywords*—Video coding, High Efficiency Video Coding (HEVC), Thor, objective video quality, subjective video quality, Rate-Distortion (RD) performance.**

I. Introduction

ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) jointly developed the High Efficiency Video Coding (HEVC) standard [1] that improves the coding efficiency of its predecessor, Advanced Video Coding (AVC) [2], by reducing the output bit-rate up to 50% for the same video quality [3]. More recently there have been initiatives, such as the Alliance for Open Media (AOMedia) led by seven big players (Microsoft, Mozilla, Google, Netflix, Amazon, Cisco, Intel Corporation), to develop a royalty free video codec and in particular Cisco just released Thor [4]. This work presents the details of the evaluation of Thor versus HEVC. The paper is organised as follows. A brief description of the main differences between HEVC and Thor at the video coding layer (VCL) is presented in Section II, and the assessment method is described in Section III. Section IV shows the objective and subjective results. Finally, conclusions are presented in Section V.

II. HEVC and Thor Differences

In this section a brief comparison between HEVC and the current version of Thor is presented, for a more detailed overview the reader can refer to [1] and [4].
Both codecs are based on the same well-known block-based hybrid approach (intra and inter prediction and transform coding) and the main differences are:

- No residual quad-tree (RQT) is used in Thor, therefore a coding block (CB) can be divided into one or at most four smaller transform blocks (TBs). The RQT technique was introduced in the HEVC standard in order to have a transform which is able to efficiently adapt to the locally varying characteristic of prediction residuals.
- No asymmetric motion partitioning (AMP) is supported for motion estimation. AMP was originally introduced in the HEVC design to improve the coding efficiency for the irregular object boundaries, however, from experiments on HEVC, the AMP tool turned out to have negligible benefits on coding performance [3].
- For intra prediction Thor uses only 8 intra prediction modes with respect to the 35 of HEVC. Of the eight modes one is the DC one, while the others are all angular ones.
- Thor adds the 64×64 DCT, but only the 16×16 low frequency coefficients are then quantised, coded and transmitted. The allowed DCT sizes in HEVC are just 8×8, 16×16 and 32×32.
- No Rate-Distortion Optimised Quantisation (RDOQ) is performed in Thor. The RDOQ is a technique which aims to improve the quantised coefficient calculation by a rough estimating the RD cost of modification or removal of the selected transform coefficient.
- For entropy coding Thor relies on context adaptive variable length code (CAVLC) whereas context adaptive binary arithmetic code (CABAC) is used in HEVC, which is a far more sophisticated technique.
- The sample adaptive offset (SAO) in HEVC is replaced in Thor by the Constrained Low Pass Filter (CLPF), which is a low pass filter applied after the deblocking filter at the super block (SB) level. In Thor the SB is the equivalent of the coding tree block (CTB) in HEVC.

III. Video Quality Evaluation

Peak Signal to Noise Ratio ($PSNR_{YUV}$) [3] as well as subjective quality evaluation results have been used for the assessment. The subjective evaluation was carried out as expert viewing according to the method described in [5], adapting it to the specific TV scenario modifying the viewing conditions as presented in [6].

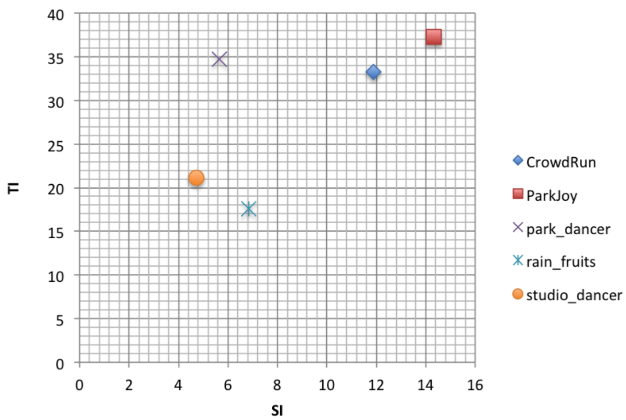


Figure 1. Spatial Information (SI) versus Temporal Information (TI) indexes of the selected contents. Both the indexes were computed in accordance with [9].

(a) CrowdRun, 1920×1080, 50 Hz.

(b) ParkJoy, 1920×1080, 50 Hz.

(c) rain_fruits, 1920×1080, 50 Hz

(d) studio_dancer 1920×1080, 50 Hz

(e) park_dancer 1920×1080, 50 Hz

Figure 2. Test sequences used for the experiment.

A. Selection of the Test Material

As suggested in [5] five different video sequences have been selected from [7] and [8], with different spatio-temporal characteristics, see Fig. 1. In Fig. 2 thumbnails of the five video sequences used for the test are shown, all of them are in 1920 × 1080 resolution at 50 fps and in YUV 4:2:0 8 bit format.

B. Configuration of Selected Codecs

HEVC reference software model HM-16.7 and the latest available version of Thor [10] were used. The Random Access (RA) configuration was set for both the encoders, with a Group of Pictures (GOP) of 8 pictures and the Intra Period of 48

TABLE I
BIT-RATE (KBPS) VALUES AND CORRESPONDING QP IN BRACKETS FOR EACH TESTED CONFIGURATION.

Sequence	**Codec**	**R1 (29)**	**R2 (32)**	**R3 (36)**	**R4 (39)**
CrowdRun	HEVC	16942.6	10764.7	6131.3	3937.7
	Thor	18145.1	10964.8	5793.5	3545.1
Sequence	**Codec**	**R1 (30)**	**R2 (33)**	**R3 (36)**	**R4 (39)**
ParkJoy	HEVC	19224.3	11574.5	7159.6	4242.9
	Thor	18523.1	11327.7	6539.8	3734.1
Sequence	**Codec**	**R1 (22)**	**R2 (27)**	**R3 (32)**	**R4 (37)**
studio_dancer	HEVC	7814.7	3879.1	1934.3	1020.0
	Thor	8565.0	4207.6	2058.8	1013.3
Sequence	**Codec**	**R1 (22)**	**R2 (27)**	**R3 (32)**	**R4 (37)**
park_dancer	HEVC	8197.7	3967.4	1883.7	889.1
	Thor	8194.6	3954.5	1822.3	808.4
Sequence	**Codec**	**R1 (22)**	**R2 (27)**	**R3 (32)**	**R4 (37)**
rain_fruits	HEVC	7921.7	3844.2	1985.5	1032.4
	Thor	8297.7	4094.5	2049.4	1027.3

pictures according to the common test conditions reported in [11]. Both the encoders were configured to use the same hierarchical B-pictures structure and the same number of reference frames, while the default quantisation parameter (QP) variations inside the GOP were used for both the codecs.

C. Test Cases

QP values and corresponding bit rates are reported in Table I. They have been chosen according to some typical broadcast scenarios.

D. Subjective Test

As discussed at the beginning of Section III , to ensure the reproducibility of the results and their validity, the method described in [5], whose Basic Test Cell (BTC) is shown in Fig. 3, was chosen. Two same halves from two different video sequences are presented simultaneously to observers by means of a split-screen presentation. In most cases one of the two halves is taken from the reference video sequence while the other from the test one. The sequence pair is presented twice in succession, once to allow scrutiny and once to allow rating. As rating scale the non-categorical SAME-DIFFERENT one [6] was selected, for which the test score is the distance between "SAME" endpoint of the scale and the mark made by the observer, expressed on a 0-100 scale. To adapt [5] to the television pictures case the laboratory for subjective video quality assessment was set up according to [6], while the distance of the subjects was set according to the design viewing distance criterion proposed in [12].

Two assessment sessions were performed to limit the number of assessors to three for each session, thus constraining the maximum observation angle and ensuring that no deviations in reproduced colour were visible to the observers. Each session was preceded by a brief training session and it was divided into four different sittings whose duration was less than half an hour. The sittings were separated from each other by 15 minutes rest period. The video trials described above were distributed across sittings by pseudo-random order, moreover at the beginning of each sitting a stabilisation phase made of dummy presentations (five for the first sitting and three for all

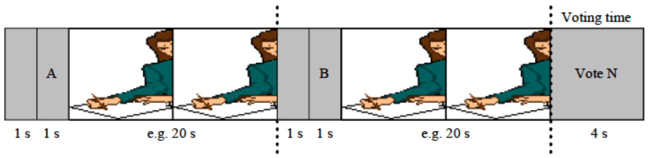


Figure 3. BTC structure.

TABLE II
BD-RATES OF THOR WITH RESPECT TO HEVC.

Sequence	DMOS BD-rate	PSNR BD-rate
CrowdRun	55.41 %	40.70 %
ParkJoy	49.91 %	36.23 %
rain_fruits	47.84 %	49.61 %
studio_dancer	34.36 %	41.58 %
park_dancer	19.73 %	23.34 %
Average	41.45 %	38.86 %

the others) was added with the aim of stabilising the observers' opinion. The monitor used to display the video stimuli was the Dolby professional reference monitor PRM-4220.

IV. Results

The subjective results expressed in terms of Difference Mean Opinion Scores (DMOS) values [13] (derived by the reference against reference presentations) are presented in the right side of Fig. 4 together with their 95% confidence intervals (CI), while in the left side the objective results are shown. To evaluate the performances of Thor and HEVC, for each sequence BD-rate has been computed [14] based on the objective and subjective ratings, as shown in Table II. HEVC encoding was taken as reference, so positive BD-rate values in Table II correspond to compression efficiency and quality losses of Thor.

As is evident from Table II HEVC far outperforms Thor for all the video sequences under test. Such difference in performance is due to the differences between HEVC and Thor presented in Section II. In particular, using the configuration files of the HM-16.7 reference software, we disabled in the HEVC encoding the following coding tools not present in Thor: RQT, AMP and RDOQ. By re-coding the same test conditions described before using this new HEVC configuration and comparing the obtained results with those previously obtained we estimated the lack of such tools in Thor with an average value of BD-rate equal to 8.59%. From such value we can assume that for the considered scenario the remaining difference in performance among Thor and HEVC, which is the main one, depends mainly on the different adopted entropy coders.

V. Conclusion and Next Steps

From the results presented above, HEVC outperforms Thor in both objective and subjective evaluations. It is likely that Thor design performance will improve if and when more sophisticated tools, not yet royalty-free, will be included, in particular for the entropy coding. However, right now Thor can not be considered a valid alternative to HEVC. Future works will cover the benchmark of Daala[1] and the inclusion of 4K video content in the sequence test set.

Acknowledgment

Computational resources were provided by HPC@POLITO, a project of Academic Computing within the Department of Control and Computer Engineering at the Politecnico di Torino (http://www.hpc.polito.it).

References

[1] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, 2012.

[2] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 560–576, 2003.

[3] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards—including high efficiency video coding (HEVC)," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1669–1684, 2012.

[4] A. Fuldseth, G. Bjontegaard, S. Midtskogen, T. Davies, and M. Zanaty, "Thor video codec," October 2015.

[5] ITU.R, "Expert viewing methods to assess the quality of systems for digital display of large screen digital imagery in theatres," *Rec. ITU-R BT.1663*, 2003.

[6] ITU-R, "Methodology for the subjective assessment of the quality of television pictures," *Rec. ITU-R BT.500*, January 2012.

[7] SVT, "The SVT high definition multi format test set," February 2006.

[8] EBU, "Technical information of ebu test sequences," March 2015.

[9] ITU-T, "Subjective video quality assessment methods for multimedia applications," *ITU-T Rec. P.910*, April 2008.

[10] "github.com/cisco/thor."

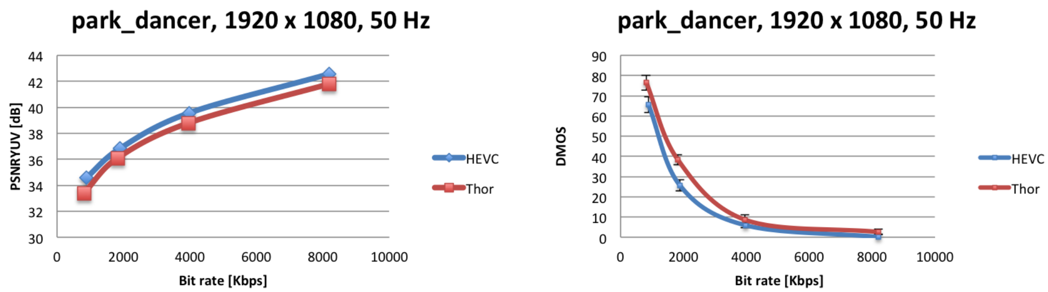
[11] F. Bossen, "Common test conditions and software reference configurations," July 2012.

[12] ITU-R, "General viewing conditions for subjective assessment of quality of sdtv and hdtv television pictures on flat panel displays," *Rec. ITU-R BT.2022*, August 2012.

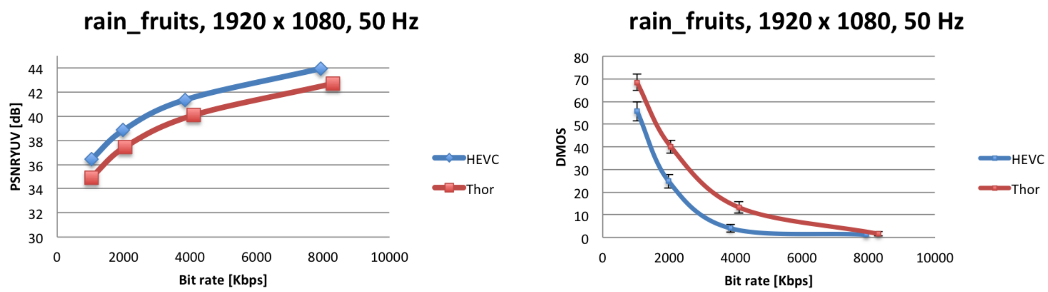
[13] VQEG, "Report on the validation of video quality models for high definition video content."

[14] G. Bjontegaard, "Calculation of average PSNR differences between RD curves," *document VCEG-M33*, Apr. 2001.

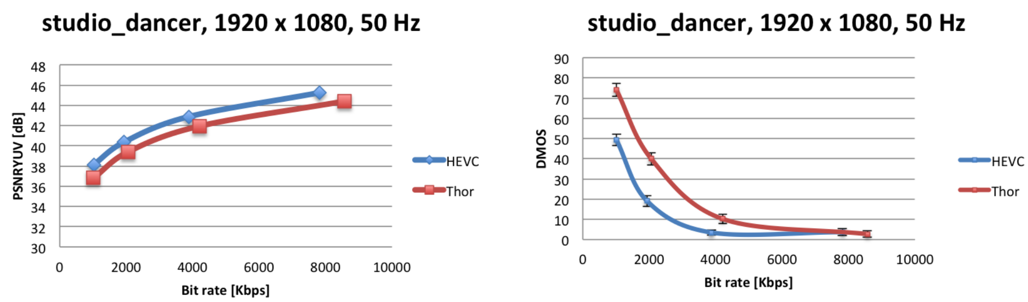
[1]Daala is a new general-purpose royalty-free video codec currently under development at Xiph.Org foundation, whose performance target is to overcome the HEVC codec.



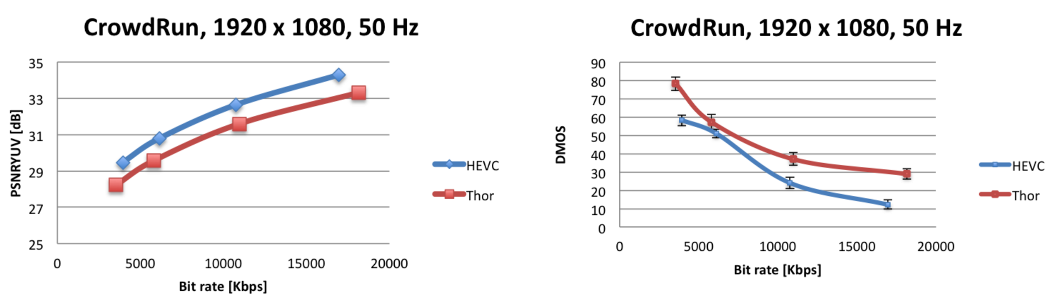
(a)



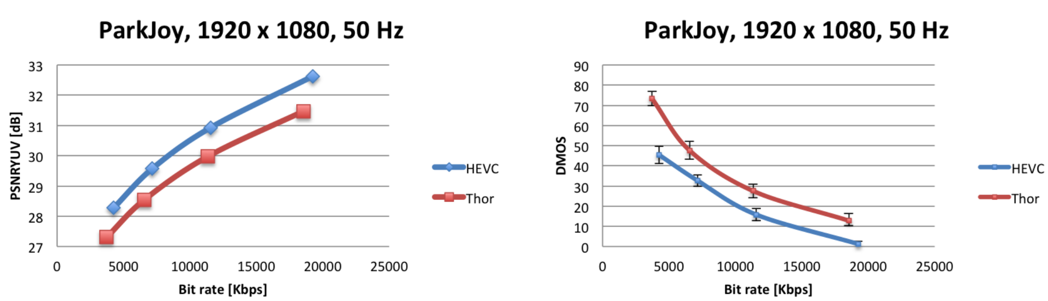
(b)



(c)



(d)



(e)

Figure 4. Objective and subjective evaluation results with the 95% confidence intervals.