Comparative Study of 8 and 10-bit HEVC Encoders

Jarno Vanne, Marko Viitanen, Ari Koivula, Timo D. Hämäläinen

Department of Pervasive Computing, Tampere University of Technology, Finland

{jarno.vanne, marko.viitanen, ari.koivula, timo.d.hamalainen}@tut.fi

***Abstract*—This paper compares the rate-distortion-complexity (RDC) characteristics of the HEVC Main 10 Profile (M10P) and Main Profile (MP) encoders. The evaluations are performed with HEVC reference encoder (HM) whose M10P and MP are benchmarked with different resolutions, frame rates, and bit depths. The reported RD results are based on bit rate differences for equal PSNR whereas complexities have been profiled with Intel VTune on Intel Core 2 processor. With our 10-bit 4K 120 fps test set, the average bit rate decrements of M10P over MP are 5.8%, 11.6%, and 12.3% in the all-intra (AI), random access (RA), and low-delay *B* (LB) configurations, respectively. Decreasing the bit depth of this test set to 8 lowers the RD gain of M10P only slightly to 5.4% (AI), 11.4% (RA), and 12.1% (LB). The similar trend continues in all our tests even though the RD gain of M10P is decreased over MP with lower resolutions and frame rates. M10P introduces no computational overhead in HM, but it is anticipated to increase complexity and double the memory usage in practical encoders. Hence, the 10-bit HEVC encoding with 8-bit input video is the most recommended option if computation and memory resources are adequate for it.**

***Index Terms*— High Efficiency Video Coding (HEVC), Video encoder, HEVC Test Model (HM), rate-distortion-complexity (RDC), Main 10 Profile (M10P)**

I. Introduction

The international video coding standard *HEVC* (*High Efficiency Video Coding*) [1], [1] is the latest milestone in the progress of video compression. It has been developed by *Joint Collaborative Team on Video Coding* (*JCT-VC*) as a joint activity of ITU-T *Video Coding Experts Group* (*VCEG*) and ISO/IEC *Moving Picture Experts Group* (*MPEG*). HEVC is published as twin text by ITU, ISO and IEC as ITU-T H.265 | ISO/IEC 23008-2. The first edition of HEVC was completed in January 2013. It includes three profiles: *Main Profile* (*MP*), *Main 10 Profile* (*M10P*), and *Main Still Picture Profile* [1].

This paper compares the *rate-distortion-complexity* (*RDC*) characteristics of the 10-bit M10P and 8-bit MP HEVC encoders whose only technical difference is the internal bit depth. Incrementing the bit depth from 8 to 10 bits enlarges the color space from $(2^8)^3$ = 16.78 million to $(2^{10})^3$ = 1.07 billion colors. Using 2^{10} shades per each of the three primary colors allows for a smoother color transition due to which the 10-bit coding format contains much less banding artifacts than the 8-bit coding format [3]. Therefore, the 10-bit coding format is adopted in *Rec. 2020 color space* [4] specification for *ultra-high definition television* (*UHDTV*). Rec. 2020 is anticipated to drive the widespread adoption of the 10-bit coding format in consumer-oriented products and services.

The RDC characteristics of the HEVC MP encoder are comprehensively evaluated by us [5] and in a couple of other previous works such as [6], [7]. However, these analyses date back to the time when M10P was not specified yet. To the best of our knowledge, the public benchmarking of the HEVC M10P and MP encoders with 10-bit test sequences is limited to a single experiment with 4K 50 fps test set [8]. With that test set, M10P is reported to get by around 5% *Bjøntegaard delta bit rate* (*BD-rate*) savings in luma component over MP. However, more comprehensive evaluations are needed before the recommended coding conditions for the 10-bit HEVC encoding can be given.

As in [5]-[8], our comparisons are conducted with the *HEVC test model* (*HM*) [9], [10]. The obtained results are based on HM 11.0 which was the latest release of HM in the beginning of our experiments. The tested configurations of HM are *all-intra* (*AI*), *random access* (*RA*), and *low-delay B* (*LB*) [11]. All tests have been carried out with a 10-bit test material that has been edited from our 10-bit 4K 120 fps test set available online [12]. Contrary to [8], our RD evaluations are not restricted to the 4K 50 fps format, but the M10P and MP encoders are compared with various resolutions and frame rates. Furthermore, our evaluations include input and output bit depth considerations for M10P and the obtained RD results are reported together with encoding complexities.

The rest of this paper is organized as follows. Section II describes our RDC analysis environment and evaluation principles. Section III compares the RDC characteristics of M10P and MP encoders as a function of different bit depths, resolutions, and frame rates. Section IV concludes the paper.

II. Analysis Setup

Fig. 1 depicts the snapshots of the 10-bit 4K 120 fps test sequences used in our experiments. They have been captured by Digiturk in May 2013 and edited by us [12]. The original video material is in 16-bit F65RAW-HFR format from which it has been converted to 10-bit 4:2:0 YUV format using FFmpeg [13]. The most essential parameters of this test set are summarized in Table 1. All sequences and their more detailed descriptions are published online in [12].

To get auxiliary test sets for our evaluations, the resolutions of these sequences have also been scaled down to 1920 × 1080 and 960 × 536 pixels. The scaling has been done with FFmpeg [13] using the bicubic algorithm. The aspect ratio of the smallest format deviates a bit from the others since its height has been cropped by four pixels after scaling (from 540 to 536 pixels) to make it divisible by eight. In addition, FFmpeg has been applied to skip intermediate frames from the sequences in order to lower their frame rates from 120 fps to 60 fps and 30 fps. Altogether, nine individual test sets have been composed of the available frame rates (120, 60, and 30

978-1-4799-6139-9/14/$31.00 ©2014 IEEE

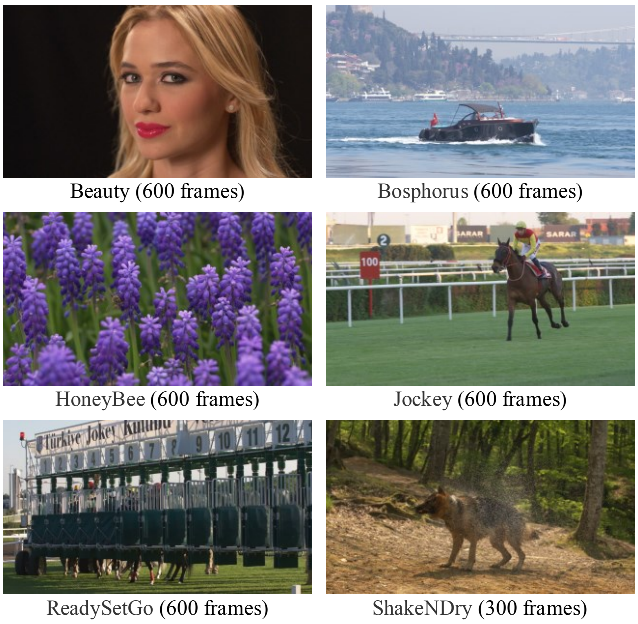

Fig. 1. Snapshots of the test sequences [12].

fps) and resolutions (3840 × 2160, 1920 × 1080 and 960 × 536). All frames of each test sequence have been encoded when benchmarking these test sets.

A. Setup for Rate-Distortion Analysis

Fig. 2 illustrates the evaluated coding schemes in which the internal, input, and output bit depths are gradually changed from 8 to 10 bits. In order to obtain a fair RD comparison, the original 10-bit YUV source video is applied as an input and the decoded output video is compared with it in all schemes.

The first scheme (*MP*) benchmarks the MP encoder for which the 10-bit source video needs to be converted to 8-bit video before encoding. The MP encoder accomplishes this pixel-wise conversion internally by rounding the pixel values with a round to nearest rule. The output HEVC stream of the MP encoder is reconstructed by the MP decoder. The RD comparison with the 10-bit source video presumes that the reconstructed 8-bit content is also extended to 10-bit video. This conversion is conducted by padding two zero LSBs in each 8-bit pixel value.

The second scheme ($M10P_{8,8}$) evaluates RD figures of the M10P encoder whose input and output are narrowed to 8 bit. To eliminate 2 LSBs in the input, the 10-bit source video is externally rounded with round to nearest rule to 8-bit data and padded back to 10-bit before encoding. The respective operation is conducted before the comparison.

The third scheme ($M10P_{8,10}$) benchmarks the M10P encoder with 8-bit input data. The encoder input is processed as in the $\mathrm{M10P}_{8,8}$ scheme but the reconstructed 10-bit output is delivered to the comparison as such without removing 2 LSBs.

The fourth scheme ($M10P_{10,10}$) evaluates the M10P encoder with 10-bit input and output data without any data conversions.

The RD comparisons between these four schemes are based on the sequence-specific bit rate differences for an identical $\mathrm{PSNR}_{\mathrm{AVG}}$ value that is a weighted average of luma (PSNR_Y)

TABLE I
PARAMETERS OF THE 10-BIT 4K 120 FPS TEST SET

Resolution	4K (3840 × 2160 pixels)
Frame rate	120 fps, progressive
Bit depth	10 bits
Subsampling	4:2:0
Color space	YCbCr
Nominal range of Y	64 - 940
Nominal ranges of Cb/Cr	64 - 960
File format	YUV

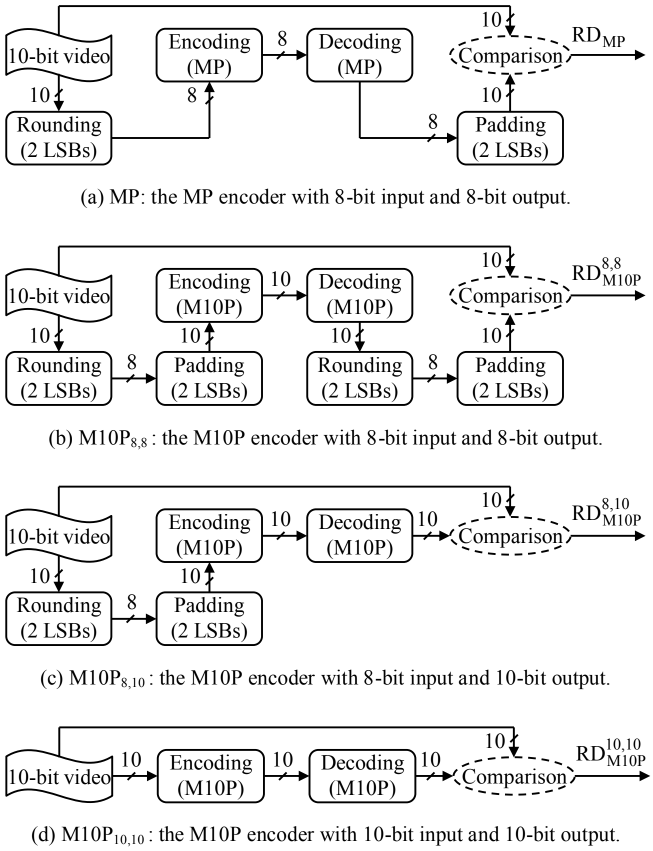

Fig. 2. The evaluated coding schemes.

and chroma (PSNR_U and PSNR_V) PSNR components [6]. For 4:2:0 color format, $\mathrm{PSNR}_{\mathrm{AVG}}$ value for the whole sequence is computed as

$$PSNR_{AVG} = (6 \times PSNR_Y + PSNR_U + PSNR_V)/8, \qquad (1)$$

where PSNR_Y, PSNR_U, and PSNR_V components are obtained by averaging their picture-specific values.

Here, the picture-specific values of PSNR_Y, PSNR_U, and PSNR_V components are each computed as

$$\mathrm{PSNR}_{\{Y|U|V\}} = 10 \times \log_{10}\left(\frac{(2^n-1)^2}{\mathrm{MSE}}\right), \qquad (2)$$

where n is the bit depth of the component and the *mean square error* (*MSE*) is computed between the original 10-bit picture and its 10-bit reconstruction.

HM uses its internal bit depth for computing (2) due to which the 8-bit HM MP encoder reports the PSNR results for the 8-bit content only. Therefore, the PSNR results of the MP encoder have been separately recomputed for the 10-bit video. Furthermore, the HM 11.0 encoder and its predecessors compute PSNR as in (2) for $n = 8$, but as

$$\mathrm{PSNR}' = 10 \times \log_{10}\left(\frac{((2^n-1)-3)^2}{\mathrm{MSE}}\right) \quad (3)$$

for $n = 10$ in order to limit the maximum value of the 10-bit coding format to 1020. This restriction makes the PSNR results of HM insensitive to padding so that the same PSNR results are yielded for the 8-bit video and its 10-bit padded extension. However, the original 10-bit values are not padded but all values (0 - 1023) are legal in practise. Therefore, the more consistent (2) is also used here with $n = 10$. The PSNR gain of (2) over (3) is

$$\mathrm{PSNR} - \mathrm{PSNR}' = 20 \times \log_{10}\left(\frac{2^n-1}{2^n-1-3}\right) \approx 0.026 \text{ dB}. \quad (4)$$

for $n = 10$. In our evaluation, this constant gain has been added to the PSNR results of the M10P schemes ($\mathrm{M10P}_{8,8}$, $\mathrm{M10P}_{8,10}$, and $\mathrm{M10P}_{10,10}$). Replacing (3) by (2) distinguishes our PSNR computation from that of [8]. In addition, our experiments specify a single BD-rate value per test sequence rather than reporting the BD-rates separately for its luma and chroma components as in [8].

The RD curves for the BD-rate computations have been interpolated through experimentally specified RD points that represent the *quantization parameters* (*QPs*) of 22, 27, 32, and 37. These QPs have been adopted from [11]. Here, the BD-rate calculations are based on the piecewise cubic interpolation since it has been shown to yield more reliable results than the 3rd order polynomial approaches [14], [15] when four RD points are used [16].

B. Setup for Complexity Analysis

Our profiling environment [5] is composed of four identical processor platforms detailed in Table 2. Only a single core per processor has been used.

The analysis relies on Intel VTune Amplifier XE profiler, which is able to report estimated cycle counts for each encoder function. This cycle-level profiling also considers the internal complexities of the functions so it is more reliable than monitoring the function calls only.

To save profiling time, all complexity results are averages of the two sequences: *Beauty* and *HoneyBee* (Fig. 1), which stand for the maximum and minimum cycle counts among the test sequences, respectively. The encoding complexities of these two sequences have been specified at the QPs of 22 and 37. In all test cases, the complexity decreases as a function of the QP, so the mean of these two corner cases gives a good approximation of the average encoding complexity.

TABLE II
PROFILING PLATFORM

Processor	Intel Core 2 Duo E8400 (2 × 3.0 GHz)
Memory	8 GB
L1 cache	2 × 32 KB (instruction) + 2 × 32 KB (data)
L2 cache	6 MB
Compiler	Microsoft Visual C++ 2010
Operating system	64-bit Microsoft Windows 7 Enterprise SP 1

III. Rate-Distortion-Complexity Analysis

Table 3 reports the sequence-specific BD-rates of our three M10P encoder schemes ($\mathrm{M10P}_{8,8}$, $\mathrm{M10P}_{8,10}$, and $\mathrm{M10P}_{10,10}$) over the MP scheme (MP) with our 4K 120 fps test set in the AI, RA, and LB cases. In the AI case, the $\mathrm{M10P}_{8,8}$, $\mathrm{M10P}_{8,10}$, and $\mathrm{M10P}_{10,10}$ schemes attain an average BD-rate gains of 1.0%, 5.4%, and 5.8% over the MP scheme, respectively. Correspondingly, the average BD-rates rise up to 4.4%, 11.4%, and 11.6% in the RA case and 5.8%, 12.1%, and 12.3% in the LB case. In all these evaluations, the RD gain of the $\mathrm{M10P}_{8,8}$ scheme fall far behind the $\mathrm{M10P}_{8,10}$ and $\mathrm{M10P}_{10,10}$ schemes, so limiting the output bit depth of the M10P encoder is not recommended. Instead, the RD gains of the $\mathrm{M10P}_{8,10}$ scheme are only 0.4, 0.2, and 0.2 percentage points lower than those of the $\mathrm{M10P}_{10,10}$ scheme in the AI, RA, and LB cases. The similar trend is visible with sequence-specific values too.

Table 4 compares the $\mathrm{M10P}_{8,10}$ and $\mathrm{M10P}_{10,10}$ schemes further by reporting their average BD-rates over the MP scheme for all our nine 10-bit test sets. Altough RD gains of the $\mathrm{M10P}_{8,10}$ and $\mathrm{M10P}_{10,10}$ schemes decrease gradually with lower resolutions and frame rates, the gap between them remains close to constant. Despite this small gap, savings in the storage and transmission capacities of a raw 8-bit input video over the 10-bit one make the $\mathrm{M10P}_{8,10}$ scheme a preferred approach. Table 4 also reports the average complexity differences between the MP and $\mathrm{M10P}_{8,10}$ schemes as *delta million cycles per frame* (*Δ Mcpf*). In each case, the complexity overhead of the M10P encoder is close to zero.

In summary, our RDC results advocate the usage of 10-bit HEVC encoding with 8-bit input video since the higher internal precision of the encoder impacts much more on the encoding result than the bit depth of the source video. However, the practical HEVC encoders tend to allow complexity optimizations that are more effective with the 8-bit encoding. This is particularly the case with hardware encoders. In addition, the impact of the bit depth on the internal memory usage cannot be benchmarked with HM 11.0 since its memory consumption remains the same in the 8-bit and 10-bit cases. The frame buffers of the practical encoders are typically byte addressable reserving one byte per 8-bit pixel but two bytes per 10-bit pixel. Hence, incrementing the bit depth from 8 to 10 doubles the sizes of these buffers. This overhead could be reduced by packing the pixels more densely in the memory, but it would violate byte addressability and limit the encoding speed due to additional alignment operations. However, introducing 8-bit input video for 10-bit encoding isolates this overhead to internal frame buffers only.

TABLE III
RD GAINS OF THE M10P ENCODER OVER THE MP ENCODER AT 4K RESOLUTION USING DIFFERENT INPUT AND OUTPUT BIT DEPTHS

Sequence	**AI**			**RA**			**LB**		
(3840×2160)	**M10P$_{8,8}$**	**M10P$_{8,10}$**	**M10P$_{10,10}$**	**M10P$_{8,8}$**	**M10P$_{8,10}$**	**M10P$_{10,10}$**	**M10P$_{8,8}$**	**M10P$_{8,10}$**	**M10P$_{10,10}$**
Beauty	-0.4%	-3.0%	-3.8%	-1.8%	-7.8%	-8.3%	-1.2%	-5.7%	-6.3%
Bosphorus	-1.2%	-6.2%	-6.4%	-4.6%	-11.2%	-11.4%	-5.6%	-11.0%	-11.3%
HoneyBee	-1.4%	-6.0%	-6.3%	-9.3%	-22.9%	-23.1%	-14.8%	-27.3%	-27.0%
Jockey	-1.3%	-9.5%	-10.1%	-4.7%	-12.9%	-12.6%	-5.5%	-14.1%	-14.6%
ReadySetGo	-0.6%	-3.6%	-3.8%	-2.4%	-5.5%	-5.8%	-3.0%	-5.8%	-6.0%
ShakeNDry	-1.1%	-4.2%	-4.5%	-3.4%	-7.8%	-8.1%	-4.5%	-8.6%	-8.8%

TABLE IV
AVERAGE RDC VALUES OF THE M10P ENCODER OVER THE MP ENCODER WITH DIFFERENT INPUT BIT DEPTHS, RESOLUTIONS AND FRAME RATES

Format	**Frame**	**AI**			**RA**			**LB**		
	rate	**M10P$_{8,10}$**	**Δ Mcpf**	**M10P$_{10,10}$**	**M10P$_{8,10}$**	**Δ Mcpf**	**M10P$_{10,10}$**	**M10P$_{8,10}$**	**Δ Mcpf**	**M10P$_{10,10}$**
3840	120	-5.4%	0%	-5.8%	-11.4%	0%	-11.6%	-12.1%	1%	-12.3%
×	60	-5.4%	-1%	-5.8%	-9.6%	0%	-9.9%	-10.8%	0%	-11.1%
2160	30	-5.4%	0%	-5.8%	-8.3%	0%	-8.4%	-9.4%	0%	-9.6%
1920	120	-3.8%	-1%	-4.2%	-6.8%	0%	-7.0%	-8.2%	0%	-8.6%
×	60	-3.8%	-1%	-4.2%	-5.2%	0%	-5.3%	-6.9%	-1%	-7.1%
1080	30	-3.8%	0%	-4.2%	-4.7%	0%	-5.1%	-5.7%	1%	-6.0%
960	120	-2.0%	-1%	-2.5%	-3.8%	-1%	-4.3%	-5.8%	0%	-6.1%
×	60	-1.9%	-1%	-2.5%	-3.2%	-1%	-3.4%	-4.3%	0%	-4.5%
536	30	-2.0%	0%	-2.5%	-2.5%	-1%	-3.0%	-3.3%	0%	-3.5%

IV. CONCLUSIONS

This paper benchmarked the RDC characteristics of the HEVC M10P encoder over MP encoder as a function of different bit depths, resolutions, and frame rates. Our RDC results advocate the M10P encoder with 8-bit input video since M10P works almost equally well for 8 and 10-bit input videos. With our 4K 120 fps test set, this scheme reduces the average BD-rate over MP by 5.4% in the AI case, 11.4% in the RA case, and 12.1% in the LB case. The respective percentages reduce with lower resolutions and frame rates being still 2.0%, 2.5%, and 3.3% with 960 × 536 30 fps test set. The encoding time of the M10P encoder equals that of the MP encoder in all our test cases. However, M10P tends to increase complexity and double the internal memory usage in practical encoders**.** To conclude, the 10-bit HEVC coding with 8-bit input video is highly recommended but only if enough computation and memory resources are available for it.

ACKNOWLEDGMENT

This work was supported in part by the European Celtic-Plus Project H2B2VS. The authors would particularly like to thank Digiturk for providing 4K test sequences.

REFERENCES

[1] *High Efficiency Video Coding*, document ITU-T Rec. H.265 and ISO/IEC 23008-2 (HEVC), ITU-T and ISO/IEC, Apr. 2013.

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

[3] A. Dueñas, A. Malamy, B. Olofsson, A. Ichigaya, S. Sakaida, S. Pejhan, L. Haglund, A. Luthra, P. Andrivon, P. Bordes, T. Jones, X. Ducloux, P. Gendron, M. Mrak, A. Cofler, J. M. Thiesse, A. Rodriguez, P. Sunna, and I. Laksono, "On a 10-bit consumer-oriented profile in High Efficiency Video Coding (HEVC)," *document JCTVC-K0109*, Shanghai, China, Oct. 2012.

[4] ITU-R Recommendation BT.2020, "Parameter values for ultra-high definition television systems for production and international programme exchange," International Telecommunication Union, Aug. 2012.

[5] J. Vanne, M. Viitanen, T. D. Hämäläinen, and A. Hallapuro, "Comparative rate-distortion-complexity analysis of HEVC and AVC video codecs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, Dec. 2012, pp. 1885-1898.

[6] 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)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, Dec. 2012, pp. 1669-1684.

[7] F. Bossen, B. Bross, K. Sühring, and D. Flynn, "HEVC complexity and implementation analysis," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, Dec. 2012, pp. 1685-1696.

[8] P. Andrivon, P. Salmon, P. Bordes, M. Arena, and P. Sunna, "Comparison of Compression Performance of HEVC Draft 10 with AVC for UHD-1 material," *document JCTVC-M0166*, Incheon, South Korea, Apr. 2013.

[9] I. K. Kim, K. McCann, K. Sugimoto, B. Bross, and W. J. Han, "High Efficiency Video Coding (HEVC) Test Model 11 (HM11) encoder description," *document JCTVC-M1002*, Incheon, South Korea, Apr. 2013.

[10] Joint Collaborative Team on Video Coding Reference Software, ver. HM 11.0 [Online]. Available: http://hevc.hhi.fraunhofer.de/

[11] D. Flynn and C. Rosewarne, "Common test conditions and software reference configurations for HEVC range extensions," *document JCTVC-L1006*, Geneva, Switzerland, Jan. 2013.

[12] Ultra video group [Online]. Available: http://ultravideo.cs.tut.fi/

[13] FFmpeg [Online]. Available: http://www.ffmpeg.org/

[14] G. Bjøntegaard, "Calculation of average PSNR differences between RD curves," *document VCEG-M33*, Austin, TX, USA, Apr. 2001, pp. 1-4.

[15] G. Bjøntegaard, "Improvements of the BD-PSNR model," *document VCEG-AI11*, Berlin, Germany, Jul. 2008.

[16] J. Wang, X. Yu, and D. He, "On BD-rate calculation," *document JCTVC-F270*, Torino, Italy, Jul. 2011.