

Central Washington University
ScholarWorks@CWU

All Faculty Scholarship for the College of the
Sciences

College of the Sciences

4-9-2017

Pre-processing techniques to improve HEVC subjective quality

D. G. Fernández

A. A. Del Barrio

Guillermo Botella

Uwe Meyer-Baese

Anke Meyer-Baese

See next page for additional authors

Follow this and additional works at: <https://digitalcommons.cwu.edu/cotsfac>



Part of the [Graphics and Human Computer Interfaces Commons](#)

Authors

D. G. Fernández, A. A. Del Barrio, Guillermo Botella, Uwe Meyer-Baese, Anke Meyer-Baese, and Christos Grecos

Pre-processing techniques to improve HEVC subjective quality

D. G. Fernández^a, A.A. Del Barrio^a, Guillermo Botella^a,
Uwe Meyer-Baese^b, Anke Meyer-Baese^c, Christos Grecos^d

^aDept. Computer Architecture, Universidad Complutense de Madrid. 28040 Madrid (Spain)
{davidgfe, abarriog, gbotella}@ucm.es

^bDept. Electrical and Computer Engineering, FAMU-FSU College of Engineering, 32310.
Tallahassee, FL (USA). umeyerbaese@fsu.edu

^cDept. Scientific Computing, Florida State University, 32306 Tallahassee, FL (USA)
ameyerbaese@fsu.edu

^dDept. Computer Science, Central Washington University, 98926-7520 Washington, WA (USA)
christos.graikos@cwu.edu

ABSTRACT

Nowadays, HEVC is the cutting edge encoding standard being the most efficient solution for transmission of video content. In this paper a subjective quality improvement based on pre-processing algorithms for homogeneous and chaotic regions detection is proposed and evaluated for low bit-rate applications at high resolutions. This goal is achieved by means of a texture classification applied to the input frames. Furthermore, these calculations help also reduce the complexity of the HEVC encoder. Therefore both the subjective quality and the HEVC performance are improved.

Keywords: HEVC, H.265, pre-processing, subjective quality.

1. INTRODUCTION

High Efficiency Video Coding (HEVC) is the next step forward in video compression being the most efficient solution for video transmission, since it consumes less than 50% of the bandwidth when compared to the previous H264 standard, but at the expense of considerably increasing the complexity of the algorithm [1-3]. HEVC, also known as H.265 [1], has been released in 2013. This new coding standardization codec is the most recent project of the Joint Collaborative Team on Video Coding (JCT-VC) between the ITU-T VCEG and ISO/IEC MPEG organizations [1]. Video standards have focused on reducing the size of the resulting bitstream. Besides this objective, the maintenance of a high level of quality has been the second factor to consider when optimizing video transmission. Because HEVC is immensely complex, several approaches that have tried to reduce its tremendous computational cost can be found in literature. It is strictly necessary to consider human perception when trying to accelerate the encoding flow, because it can lead to a subjective quality improvement achieving a better user experience. The rest of the article is organized as follows. Section 2 discusses the state of the art and Section 3 describes the proposal in detail. The experiments and the performance of the proposed algorithm are discussed in Section 4. Our conclusions and lines for future studies are given in Section 5.

2. RELATED WORK

In contrast to previous video coding standards, HEVC uses a flexible quad-tree coding block partitioning structure that enables the use of large and multiple sizes of coding, prediction, and transform blocks. This system is more efficient but also more complex, providing the encoder the capacity to select the proper partition sizes [1-2].

It is widely accepted that image blocks in which there is a spatially predominant direction should not be divided into smaller blocks in order to reduce the associated rate distortion (RD) cost [4-10]. Several gradient-based algorithms apply this concept to reduce the HEVC encoder complexity [10-15]. A similar idea can be employed to areas with homogenous textures also known as smooth regions [4-6][8][11-13][16-19]. In our proposal, an optimized CU size decision algorithm is employed on smooth regions to accelerate the encoding process based on three of our previous works [4-6]. The

studies presented in [4], [5] and [6] apply an optimized CU size decision after a smooth region classification, if a CU is classified as smooth the quad-tree recursive process is stopped. Moreover, in the work presented in [4] proved that the smooth-based detection and optimization can coexist with other optimizations potentially applicable to the rest of the image, such as gradient-based algorithms [7]. In [4] the combined CU size decision method using both algorithms is implemented on the x265 software [20]. The x265 software is a H265/HEVC video encoder application library that allows HD real-time encoding. This software is a fast and computationally efficient implementation available under the terms of the open source GNU GPL 2 license [20].

In [5] the proposed CU size decision method was customized for lossless compression mode [2], which is quite usual in some medical applications to avoid coding artifacts that could interfere with diagnosis. In that article all the tests targeted medical images. In [6] an optimized CU size decision algorithm is proposed to reduce the computational cost of quad-tree partitioning by means of spatial and temporal homogeneity analysis and classification, which are directly applied to the input image. Furthermore, this image pre-analysis is performed using logic units and embedded hardware on a GPU, thus avoiding unnecessary waiting states, so the computational cost associated with this process is zero for the processor in charge of the encoding process.

It is also well known that pre-processing algorithms can be used to improve the video compression efficiency by means of noise reduction and the elimination of artefacts [21]. Video pre-processing algorithms are not standardized, so they could be an important factor for raising a video encoder above the others. According to the human visual system (HVS), textured areas can be coded with a higher amount of noise, before the noise becomes noticeable [22]. Furthermore, HVS is more sensitive to details in the areas with homogeneous texture activities, so it is annoying to observe artefacts in these areas such as the sky or a blank wall. Authors in [22] propose introducing distortion by varying the value of the Quantification Parameter (QP) in textured areas, which can be coded with a higher amount of noise, that is, before the noise becomes noticeable. On the contrary, in our proposal, we have decided to improve the quality of the regions with homogeneous textures to obtain a better global perception from the HVS perspective with a simpler texture classification. After this classification is conducted, a filter is applied to regions with a chaotic distribution to reduce the amount of detail that the encoder needs to process. This filter can be applied without a negative effect because chaotic regions have a lot of high frequency coefficients that are imperceptible by the HVS.

Therefore, two algorithms are proposed to improve subjective quality in the HEVC encoder in this article. The first one is based on the detection of regions with homogeneous texture. This classification produces a binary map which is sent to the encoder in order to improve the quality of these regions by reducing the quantization parameter (QP). The second algorithm identifies areas with a chaotic distribution using the outcome of the previous classification as well as a gradient-based algorithm. If an area is classified as a part of a chaotic distribution, then a median filter is applied to replace its pixels. The pre-processed frames are then handed to the encoder. Finally, a better global perception from the HVS perspective is obtained after the HEVC encoding.

3. IMPROVING SUBJECTIVE QUALITY

3.1 Homogeneous Regions

The detection of homogenous regions is based on the algorithm proposed in [4], [5] and [6] to reduce the complexity of the coding unit (CU) size decision process. This detection is based on the relationship between DC and AC coefficients. These sets of coefficients are obtained after applying a Discrete Cosine Transform (DCT) within the video encoder. The DC coefficient is the one with zero frequency, while the AC coefficients are the remaining ones. The main drawback of this calculation in the frequency domain is its high computational cost. Nonetheless, according to the Parseval's relation [23], the time and frequency domains are equivalent representations of the same signal. Therefore, it is possible to work in the pixel domain to obtain a DC ratio (DC_R) as it is explained in [4], [5] and [6].

The DC ratio of each block is compared with a threshold (DC_{TH}), and if $DC_R > DC_{TH}$ the block is classified as homogeneous. The output of this analysis is a binary map specifying whether or not a region has been categorized as homogeneous. This map is sent to the encoder. Inside the encoding loop if the image block is classified as homogeneous, a QP decrement (QP_D) is applied to the QP value given by the rate control (QP_R). The rate control algorithm dynamically adjusts the quantization parameter to achieve a target bitrate. High QP values provide poor quality and few bits. Low QP values produce encodings with better quality and more bits. Therefore, decreasing QP provides a quality improvement.

Thus the QP value for homogeneous regions will be $QP_H = QP_R - QP_D$. Because of this QP modification, the number of bits generated will be higher, so the rate control will decide to increase QP_R to compensate for that variation in the next coding units. Hence, our algorithm must be careful when selecting QP_D , as a noticeable decrement might cause the rate control to raise QP_R too much, diminishing the overall image quality. Therefore, the value of QP_D must provide an increase in subjective quality, but without negatively affecting QP_R . Thus, the decrease introduced by QP_D depends on the current QP_R value, as Figure 1 shows. An example of QP map is shown in Figure 2. In this Figure, the QP map has been obtained after applying the proposal in the *in_to_tree* sequence encoded at 512Kbps.

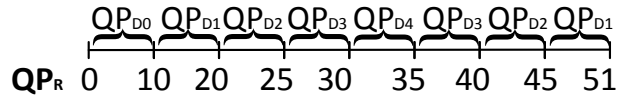
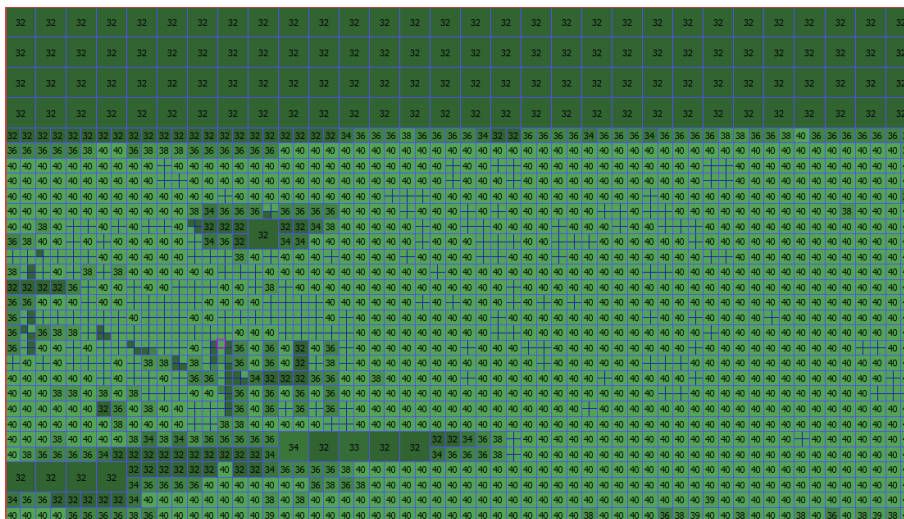


Fig. 1. QP_D zones for different QP_R values.



a)



b)

Fig. 2. a) Encoded Frame b) QP distribution after applying the proposal.

As can be seen in Figure 2, the smooth regions as the sky and the road have been encoded using a QP_H value ranging from 32 to 38 which are lower than the QP_R (40) employed in the rest of the image.

3.2 Chaotic regions

Since homogeneous areas possess low complexity, reducing QP value will theoretically cause a negligible increase in the number of generated bits. Nonetheless, the bitstream increase is not affordable for some sequences at low bitrates. In order to mitigate this increase, a median filter is applied to areas with a chaotic texture distribution to eliminate noise and imperceptible information as the removal of unnecessary visual information delivers equivalent perceived quality while reducing bit rates. The chaotic region detection is applied to every pixel in the image. The procedure is based on two simple conditions. First, the pixel must not belong to a homogeneous block. The second condition involves comparing the gradient amplitude (A_{GR}) obtained for this pixel with a predetermined threshold (GR_{TH}). If $A_{GR} < GR_{TH}$, then the pixel is within a chaotic region and is replaced by the output of a 3x3 median filter. Since this type of filter could introduce blurring in the image, it is important to determine the suitable threshold. Finally, it must be noted that we decided to use the gradient amplitude for this detection because it is a calculation used in several works [10-15] to reduce the HEVC encoder complexity. Nevertheless, neither of these works considers subjective quality improvements. Figure 3 shows the differences between the original and filtered image for the first frame of the *flower garden* sequence which is accessible for free download from [24]. The differences are highlighted using blue and red colors. Red pixels indicate a big difference between the luminance values.

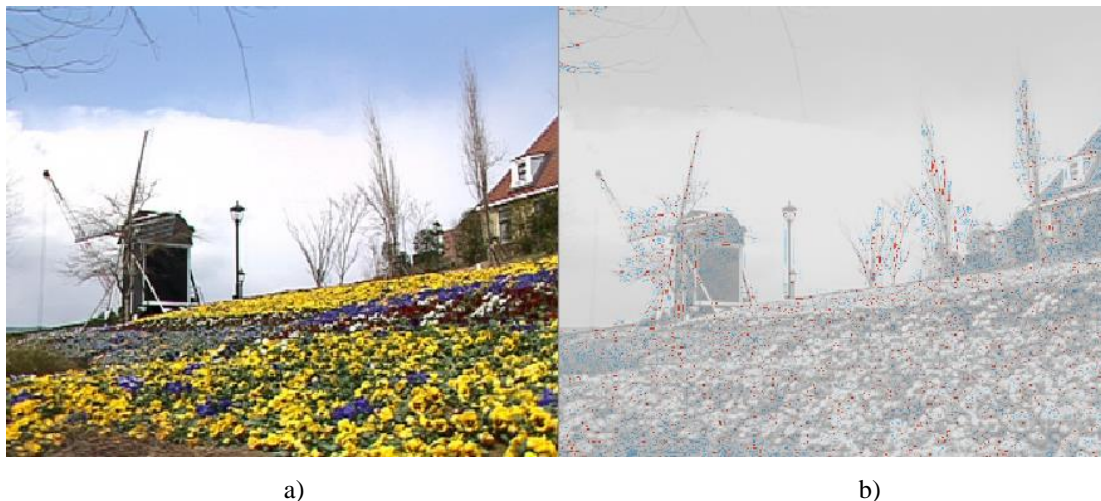


Fig. 3. Comparison between original and filtered image after applying the chaotic detection.
a) Original Frame
b) Difference between original and filtered image (3x3 Median filter applied to chaotic regions).

4. EXPERIMENTS

The proposal has been tested by encoding several sequences with different features. The corresponding modifications have been performed on the x265 [20] using the ultrafast preset [20], but they could be easily implemented in any encoder regardless of the hardware architecture. The tests were performed using an Intel Core i5-4460@3.40GHz microprocessor, implementing the texture analysis and filtering process in its embedded GPU (Intel HD Graphics 4600) using OpenCL v1.1 [25]. These processes take 7.67 ms on average for one 1080p frame and 3.2 ms for a 720p frame. This is done using only the input images and introducing one frame of delay in the encoding process to avoid stalls for GPU data. The quality comparison has been performed based on SSIMb, since this metric is well-correlated with the



a)



b)

Fig. 4. Subjective Quality Comparison for *in_to_tree* sequence (1080p)

a) Decoded image without perceptual improvement.

b) Decoded image applying the proposal

human perceptual experience for video compressed sequences [26]. Figure 4 shows how the proposal significantly reduces unpleasant effects such as blurring, blocking, mosquito's noise and banding for frame 240 in the *in_to_tree* sequence encoded at 512Kbps. In this frame the SSIMb difference is 0.09 (11.5% improvement).

The bitstream size saving allows achieving better quality while keeping the same bitrate as can be seen in Table I. The SSIMb difference has been measured with respect to the baseline case without our modifications for different bit-rates in percentage terms. Table I shows the average results for the HD sequences specified in [27] at low bit rates. The document presented in [27] is very useful for performance comparisons because it establishes a set of common test conditions for this purpose. Table I also shows how the maximum quality improvements are noticeably higher than the average values, which provides a better global perception from the HVS perspective as can be observed using the Mean

Opinion Score (MOS) metric. This subjective comparison has been performed based on ITU-R BT.500 [28]. Specifically, the Double Stimulus Impairment Scale (DSIS) method [28] has been applied in the tests. In this method videos are shown sequentially in pairs: the first one is the uncompressed original video, whereas the second one is the encoded version. After playing these videos, 20 experts were asked to give their opinion using an impairment scale: 5- Imperceptible; 4-Perceptible, but not annoying; 3-slightly annoying; 2-Annoying; 1-Very annoying.

Table I. Quality comparison based on SSIMb and MOS metrics

Sequences	BitRate	Average Δ SSIMb (%)	Maximum Δ SSIMb (%)	MOS Ref	MOS Our
Class B (1080p)	512Kbps	2.34	12.03	2	2.72
	750Kbps	2.12	10.07	2.44	3.18
	1Mbps	2.01	9.11	3	3.56
	2Mbps	1.5	9.01	3.4	4.1
	3Mbps	1.06	6.01	3.98	4.46
Class E (720p)	512Kbps	2.21	11.22	3	3.3
	750Kbps	1.94	8.71	3	4
	1Mbps	1.34	7.14	4	4.35
	2Mbps	1.23	7.27	4	4.23
	3Mbps	0.076	4.34	4.65	4.8

A detailed example of the SSIMb values obtained after applying the proposal can be seen in Figure 5 for the *BasketballDrive* sequence at 750Kbps.

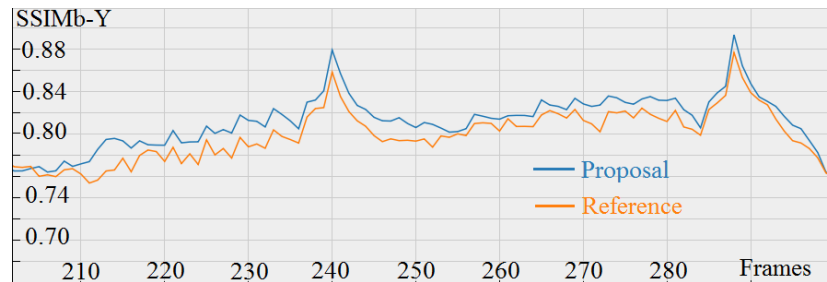


Fig. 5. SSIMb values at the middle of the *BasketballDrive* sequence.

4.1 Performance

This section proves how the calculations used to improve the subjective quality can also help reduce the HEVC encoder complexity. For that purpose the algorithm described in [4-6] has been used to reduce the HEVC encoder complexity in spatially homogeneous regions, and the gradient-based algorithm proposed in [7] to the rest of the image. The formula that defines the encoding time measurement is depicted by equation 1:

$$\Delta\text{Time} (\%) = \frac{\text{Proposal Time} - \text{Reference Time}}{\text{Reference Time}} * 100 \tag{1}$$

As can be seen in Table II the encoding time reduction is above 25% (36.66 % in the best case) at the cost of a negligible Bjontegaard delta bit-rate (BD-BR) increment [29]. It should be noted that a low BD-Rate percentage increase implies better coding efficiency. The HM16.2 test model [2] was used for these experiments under the conditions described in [27] for the *All Intra-Main* configuration, since it could be used for comparison purposes.

Table II. Integration with existing works [4][7]

Sequence	BD_BR (%)	Δ Encoding Time (%)
Class B (1080p)	0.8	-25,38
Class E (720p)	0.8	-36,67

5. CONCLUSION

In this paper a subjective quality improvement has been proposed based on the Human Visual System. This has been done by utilizing some pre-processing algorithms for detecting chaotic and homogeneous regions, which are also useful to reduce the HEVC complexity. Results support our claims for several bitrates, increasing subjective quality while reducing the encoding time.

6. ACKNOWLEDGMENTS

This work has been partially supported by Spanish research Projects TIN 2015-65277-R and TIN-2012-32180, as well as the UCM-Banco Santander Grant PR26-16/20B-1.

7. REFERENCES

- [1] G. Sullivan, J. 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, pp. 1649–1668, Dec 2012.
- [2] K. McCann, C. Rosewarne, B. Bross, M. Naccari, K. Sharman, and G. Sullivan, "High Efficiency Video Coding (HEVC) encoder description 0v16 (HM16)," *JCT-VC High Efficiency Video Coding N14 703*, 2014.
- [3] H. Koumaras, M. Kourtis, and D. Martakos, "Benchmarking the encoding efficiency of h.265/hevc and h.264/avc," in *Future Network Mobile Summit (FutureNetw)*, 2012, July 2012, pp.
- [4] D. G. Fernandez, A. A. Del Barrio, G. Botella and C. Garcia, "4K-based intra and inter prediction techniques for HEVC", *Proc. SPIE, Real-Time Image and Video Processing 2016*, vol. 98970B, April 2016.
- [5] D. G. Fernández, A. A. Del Barrio, G. Botella, C. García, U. Meyer-Baese, A. Meyer-Baese, "HEVC optimizations for medical environments", in *Proc. SPIE 9871, Sensing and Analysis Technologies for Biomedical and Cognitive Applications 2016*, 98710B, May 19, 2016.
- [6] D. G. Fernández, A. A. Del Barrio, G Botella and C. García. "Fast and effective CU size decision based on spatial and temporal homogeneity detection". *Multimedia Tools and Applications*, 1-21.
- [7] S. Na, W. Lee, and K. Yoo, 'Edge-based fast mode decision algorithm for intra prediction in hevc', in *IEEE International Conference on Consumer Electronics (ICCE)*, Jan 2014, pp. 11–14, doi: 10.1109/ICCE.2014.6775887
- [8] G. He, D. Zhou, and S. Goto, "Transform-based fast mode and depth decision algorithm for hevc intra prediction," *IEEE 10th International Conference on ASIC (ASICON)*, Oct 2013, pp. 1–4.
- [9] T. Mallikarachchi, A. Fernando, and H. Arachchi, "Efficient coding unit size selection based on texture analysis for hevc intra prediction," *IEEE International Conference on Multimedia and Expo (ICME)*, July 2014, pp. 1–6.
- [10] S. Na, W. Lee, and K. Yoo, "Edge-based fast mode decision algorithm for intra prediction in hevc", *IEEE International Conference on Consumer Electronics (ICCE)*, Jan 2014, pp. 11–14.
- [11] Y. Zhang, Z. Li, and B. Li, "Gradient-based fast decision for intra prediction in hevc," in *Visual Communications and Image Processing (VCIP)*, Nov 2012, pp. 1–6.
- [12] Y.-C. Ting and T.-S. Chang, "Gradient-based pu size selection for hevc intra prediction," *IEEE International Symposium on Circuits and Systems (ISCAS)*, June 2014, pp. 1929–1932
- [13] H. Zhang, and Z. Ma, "Fast intra mode and CU size decision for HEVC". *IEEE Transactions on Circuits Systems for Video Technology*, vol.PP, iss.99, Apr. 2016.
- [14] M. U. K. Khan, M. Shafique, M. Grellert, and J. Henkel, "Hardware-software collaborative complexity reduction scheme for the emerging hevc intra encoder," *Design, Automation Test in Europe Conference Exhibition (DATE)*, 2013, March 2013, pp. 125–128.
- [15] T. Ye, D. Zhang, F. Dai, and Y. Zhang, "Fast mode decision algorithm for intra prediction in hevc," in *Proceedings of the Fifth International Conference on Internet Multimedia Computing and Service*, 2013, pp.300–304.
- [16] G. Tian and S. Goto, "Content adaptive prediction unit size decision algorithm for hevc intra coding," in *Picture Coding Symposium (PCS)*, May 2012, pp. 405–408.

- [17] B. Min and R. Cheung, "A fast cu size decision algorithm for hevc intra encoder," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. PP, no. 99, pp. 1–1, 2014.
- [18] L. Shen, Z. Zhang, and Z. Liu, "Effective CU size decision for HEVC intra coding," *IEEE Transactions on Image Processing*, vol. 23, no. 10, October 2014.
- [19] X. Liu, Y. Liu, P. Wang, C. Lai, and H. Chao "An adaptive mode Decision algorithm based on video texture characteristics for HEVC intra prediction". *IEEE Transactions on Circuits Systems for Video Technology*, vol.PP, iss.99, Apr. 2016.
- [20] x265 project, <http://x265.org/>, accessed January 2017.
- [21] Yves Faroudja, "Improving Video Streaming and File Compression Efficiency without Affecting Quality" in *SMPTE 2014 Annual Technical Conference & Exhibition*.
- [22] K. Minoo and T. Nguyen, 'Perceptual video coding with h.264,' in *Asilomar Conference on Signals, Systems and Computers*, Oct 2005, pp. 741–745, doi: 10.1109/ACSSC.2005.1599851
- [23] S. W. Smith, 'The Scientist and Engineer's Guide to Digital Signal Processing', (California Technical Publishing, 1997, second edition)
- [24] xiph.org, Derf's test media collection, <https://media.xiph.org/video/derf/>, 2017.
- [25] Khronos OpenCL Working Group. "The OpenCL Specification Version 1.1". Revision 44.
- [26] Rahul Gaurav and Hasan F. Ates. 'Efficient Quality of Multimedia Experience Using Perceptual Quality Metrics' in *Conference on Advanced Computing, Networking and Informatics*, 2015, doi: 10.1007/978-81-322-2538-6_50
- [27] F. Bossen, 'Common test conditions and software reference configurations,' *JCT-VC Document*, JCTVC-K1100, October 2012.
- [28] ITU-R BT.500 "Methodology for the subjective assessment of the quality of television pictures", *International Telecommunication Union*, Geneva (2012).
- [29] G. Bjontegarrd, 'Calculation of average PSNR differences between RD curves,' *ITU-T SC16/Q6 13th VCEG meeting*, Austin, Apr. 2001.