Accepted in the 2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE 2016), 2016

RDO Cost Modeling for Low-Complexity HEVC Intra Coding

Mohammadreza Jamali and Stéphane Coulombe
École de technologie supérieure, Université du Québec, Montreal, Canada
mohammadreza.jamali.1@ens.etsmtl.ca, stephane.coulombe@etsmtl.ca

***Abstract*—High efficiency video coding (HEVC) is the newest international standard for video compression, providing improved coding performance that achieves compression ratios up to 50% higher than those obtained with H.264/AVC. However, this improvement comes at the expense of high computational complexity and coding time. In this paper, we propose a novel method for fast and low-complexity intra HEVC mode decision based on rate-distortion optimization (RDO) cost modeling, which permits the exclusion of non-promising candidates from the RDO processing. To achieve even more complexity reduction, an additional rough most probable modes examination is coupled with the main algorithm. Experimental results show that the proposed algorithms reduce the encoding time by 41.8% on average, with a negligible quality loss of 0.058 dB (BD-PSNR) for *all-intra* scenarios, as compared to the HEVC reference implementation, the HM 15.0.**

Keywords—HEVC; video coding; intra coding; mode decision; rate-distortion optimization

I. INTRODUCTION

High efficiency video coding (HEVC) [1] is the most recent video coding standard developed by the joint collaborative team on video coding (JCT-VC). HEVC doubles the compression ratio compared to H.264/AVC, at the same video quality. This improved performance is gained by introducing several new coding tools. With intra coding, while H.264/AVC considers a maximum of 9 prediction modes, HEVC employs 35 modes allowing the encoder to efficiently exploit the spatial correlation in a frame by performing more precise angular pixel prediction. Fig. 1 shows these intra prediction modes. However, the significant coding improvement of HEVC requires increased computational complexity at the encoder. For intra coding, this is mainly due to computing the highly-complex rate-distortion optimization (RDO) for all modes. In view of this, fast intra algorithms are highly desirable to reduce the encoding time, computational complexity and energy consumption.

Several research studies have been conducted to address the problem of the fast intra mode decision. To exploit the spatial correlation in a frame, neighboring blocks' modes have been used as candidates for the best mode of the current block [2]. However, this approach results in a domino effect, i.e., a wrongly decided mode may be propagated, leading to a significantly reduced visual quality. In [3], the properties of

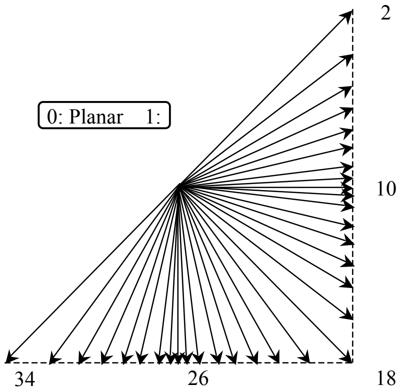

Fig. 1. HEVC intra prediction modes

the neighboring reference samples were employed to determine the number of modes that need to be processed by RDO. Other approaches have used the content of the block, like the dominant directions (edges), to specify the best intra mode [4-5]. These algorithms are very effective in excluding irrelevant directional modes. However, they cannot accurately predict the best intra mode since the objects' edges are not exactly aligned with the directions of HEVC intra modes, making such predictions difficult. In [6], the authors evaluate modes using a low-complexity measure, such as the sum of the absolute transformed differences (SATD), and compute the RDO cost only on a number of selected candidates having the lowest SATDs. However, time reduction is limited, since a predefined number of candidates are considered.

In this paper, we propose a novel method to exclude non-promising prediction modes from further RDO processing by applying a Gaussian model for the RDO cost. This modeling results in a very low-complexity and high quality intra mode decision approach since it can effectively identify the smallest possible number of intra modes that need to be investigated by rate-distortion optimization. It uses a low-complexity cost based on SATD. Other works that have used this metric keep the N best modes in SATD and evaluate them in RDO. However, they consider neither the actual SATD values nor the relationship between SATD values and the RDO costs in establishing the list of the most promising modes. Using the N best modes, they either include candidates that have no potential of winning, and waste computational resources, or select an insufficient number of candidates, and reduce the visual quality. Conversely, our method carefully selects the

This work was funded by Vantrix Corporation and by the Natural Sciences and Engineering Research Council of Canada under the Collaborative Research and Development Program (NSERC-CRD 428942-11).

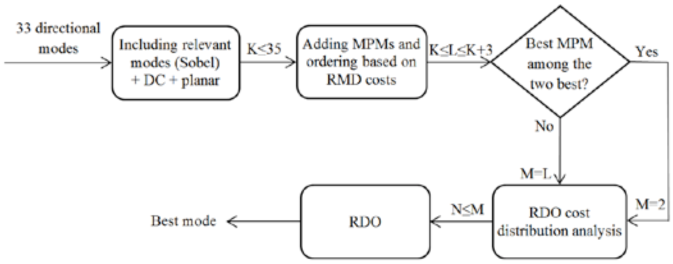

Fig. 2. Block diagram of the proposed method

candidates based on individual SATD values. Additionally, we added a rough most probable modes examination phase to our RDO modeling, which eliminates some modes when it is clear that the current block follows a similar prediction pattern as the neighboring blocks. This technique provides extra complexity reduction, and results in a faster encoder. The accuracy of these algorithms is high enough that the reconstructed video quality is hardly affected.

This paper is organized as follows. In section II, we introduce the mode decision mechanism of the HEVC intra coding. Section III presents the different stages of the proposed method. Experimental results are shown in section IV, and finally, section V concludes the paper.

II. HEVC Intra Mode Decision

HEVC applies intra coding to exploit spatial correlations inside a frame. Intra coding involves mode decision, which ideally consists in finding the best intra prediction mode, i.e., the mode among 35 modes that results in the best rate-distortion performance. These modes include 33 angular, DC and planar modes. In the HM [7], which is the HEVC test model, the best mode is selected by an RDO process, and is the mode with the lowest RDO cost. This optimization provides a tradeoff between the reconstructed video quality and the number of bits used for encoding. RDO is based on the RDO cost ($Cost_{RDO}$), which is defined as [8]:

$$Cost_{RDO} = (SSE_{luma} + \omega_{chroma} \times SSE_{chroma}) + \lambda_{Mode} \times R_{Mode} \quad (1)$$

where SSE (sum of squared errors) is a measure of distortion between the original and reconstructed blocks, and must be evaluated for both *luma* and *chroma* components, ω_{chroma} is a weight that is a function of the quantization parameter (QP), λ_{Mode} is the Lagrange multiplier, and R_{Mode} is the total number of bits for coding the block. $Cost_{RDO}$ needs to be evaluated for all 35 modes to find the optimum intra mode. It should be noted that RDO is computationally very expensive, and slows down the HEVC encoder. To reduce the number of RDO computations required to select the best mode, HM may alternatively use the rough mode decision (RMD) process, where it selects the N modes with the lowest RMD costs and runs the RDO only for these modes. The RMD cost ($Cost_{RMD}$) for a block is defined as [8]:

$$Cost_{RMD} = SATD_{luma} + \lambda_{pred} \times R_p . \quad (2)$$

$SATD$ is computed between the original luma block and its prediction block, which is obtained using one of the intra modes, and represents the distortion between these two blocks. R_p is the number of bits used to signal the selected mode and λ_{pred} determined using:

$$\lambda_{pred} = \sqrt{\lambda_{Mode}} . \quad (3)$$

In the RMD process N is equal to 8, 8, 3, 3 and 3 for 4×4, 8×8, 16×16, 32×32 and 64×64 block sizes, respectively. The problem with this approach is that it requires the use of a fixed number N for all video classes, QPs etc. This results in quality degradation when the best mode is outside the set of these N modes, or leads to unnecessary computations when fewer than N modes are needed to be considered in the RDO process. To solve this problem, we propose, in the subsequent section, an adaptive method based on RMD cost, which can efficiently reduce coding complexity without significant loss of quality.

III. Proposed Method

In this section, we propose a method for classifying the modes into *promising* and *non-promising* modes. The highly complex RDO cost computation is performed only for the promising modes. Fig. 2 illustrates the block diagram of the proposed method. Its various stages are presented in the subsequent sub-sections. Since the edge detection algorithms are less complex than computing RMD costs, edge detection is used to include the relevant angular candidates.

A. Including Relevant Angular Modes

Our method for fast and low-complexity HEVC intra coding is a multistage approach, which includes a previously proposed edge detection process and two novel techniques proposed in this paper. The first stage intends to include only some relevant angular modes using a low-complexity edge detector proposed in [9]. Since the edge detection algorithm does not consider DC and planar modes, we add these two modes to our candidates to form a list of K modes ($K \leq 35$). K is configurable, and can be adjusted to provide a tradeoff between complexity and quality.

B. Rough Most Probable Modes Examination and Ordering

To take into account the spatial correlation between blocks, we consider modes of the neighboring blocks as potential candidates in order to find the best mode of the current block. We use this concept in our work to include the neighboring modes if they were not included by the edge detection algorithm. In this stage, the list of candidates is first augmented by three most probable modes (MPMs) from neighboring blocks [10], resulting in L candidates with $K \leq L \leq K+3$, depending on whether or not each of the most probable modes already included in the list. We call these most probable modes m_{mpm1}, m_{mpm2} and m_{mpm3} and we define $m_{mpmBest}$ as:

$$m_{mpmBest} = \underset{m \in MPMs}{\arg\min}\, Cost_{RMD} . \quad (4)$$

After ordering the candidates selected by the edge detection and the MPMs, based on the RMD cost, we check whether the $m_{mpmBest}$ is one of the two best modes. If that is the case, only the two modes with the lowest RMD cost are considered for next step. Otherwise, all the selected candidates are considered by RDO cost distribution analysis. The number of candidates, after this stage, should be M, while M is either L or 2.

C. Statistical RDO Cost Modeling Based on RMD Cost

In this sub-section, we develop a model for the RDO cost, based on the RMD cost, to avoid performing RDO for all candidates. Given that there a relationship exists between RMD and RDO costs, and that the RMD process is much less complex than that of the RDO, we can determine, to some degree, the order of RDO costs based on the RMD costs. Since obtaining this order may not be perfect, the mode with the lowest RMD cost is not necessarily the same as the mode with the lowest RDO cost, i.e., the best mode. Thus, we adaptively select N modes with the lowest RMD costs such that there is a high probability that the best mode will lie within the set of selected modes.

According to our observations on the $Cost_{RDO}$ distributions, for blocks having specific $Cost_{RMD}$ values, the $Cost_{RDO}$ fits some well-known probability distributions very well. In other words, there is a statistical relationship between RMD and RDO costs. In view of this, for each $Cost_{RMD}$ value range, we assign a probability density function for $Cost_{RDO}$. Thus, we remove the modes having $Cost_{RMD}$ values with a very small probability of having the best $Cost_{RDO}$. We divide the entire range of RMD costs into multiple bins for which the associated RDO costs histogram is obtained. We then examine how closely a number of known distributions can best fit the obtained histogram. Fig. 3 shows the histogram of the empirical data estimated by four of the distributions for the *RaceHorses* video sequence for an 8×8 block size and QP of 32. $Cost_{RMD}$ ranges from 2903 to 2945 for this figure. It can be seen that the normal distribution provides an accurate fit to the RDO cost histogram. Similar results are also observed for other video sequences, block sizes, QPs and $Cost_{RMD}$ values.

In the light of these results, the distribution of RDO cost is modeled by the normal distribution that lets us select the most promising modes based on RMD costs. To apply this modeling to our problem, we need to have normal fitted distributions for RDO costs of different bins of the RMD cost range. Fig. 4 shows an example of these distributions associated with candidate modes. Again, this figure is achieved for *RaceHorses* with 8×8 block sizes and the QP is set to 32. Similar results are obtained for other sequences, block sizes and quantization parameters. We adaptively select N distributions among M distributions related to M candidates. Normally, if all intra modes are considered, M is equal to 35. But if the two approaches presented in sub-sections III.A and III.B were applied, M would be less than 35 and there would be no need to compute the RMD cost for the excluded angular modes. This results in additional complexity reduction compared to the case of the exhaustive approach, since edge detection computation is less complex than RMD cost computation. The set of selected candidates, before applying

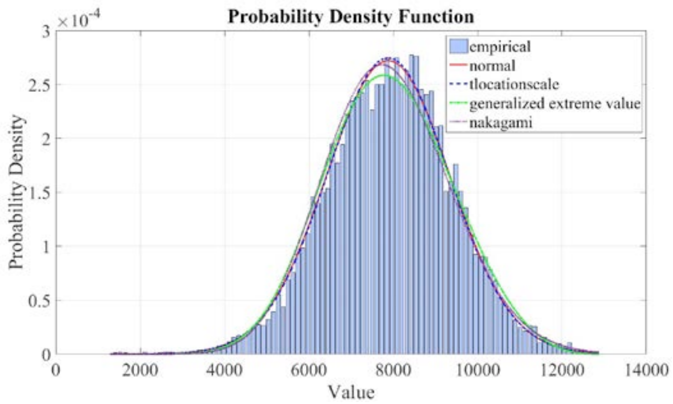

Fig. 3. RDO cost distribution for *RaceHorses* and size 8×8 blocks, QP=32 and $Cost_{RMD}$ between 2903 and 2945

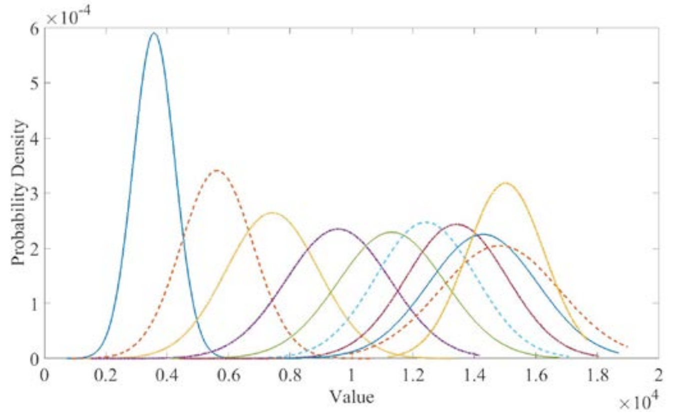

Fig. 4. RDO cost distributions for different bins of RMD cost, *RaceHorses*, size 8×8 blocks, QP=32

RDO cost distribution analysis, is denoted by ψ, defined as:

ψ = *{DC+ Planar+ Selected modes based on edge detection+ MPMs}* = $\{m_1, m_2, \ldots m_M\}$.

As stated earlier, ψ would have only two members if the condition mentioned in sub-section III.B were satisfied. Knowing the RMD costs for M modes, bins associated with these costs are determined. Based on the model, for each bin, there is an RDO cost distribution; thus we have M normal overlapping distributions, i.e., $X_i \sim N(\mu_i, \sigma_i^2)$ each associated with a candidate mode. We consider the mode associated with the distribution with the lowest mean as the temporary best mode (m_{RMDmin}), and compare it with the others to classify them as promising or non-promising candidates. At the end of the process, promising modes form a set called P, defined as:

P = *{Promising modes based on RDO cost distribution analysis}* = $\{m_1, \ldots m_N\}$.

where $P \subseteq \psi$ and $N \leq M$. m_{RMDmin} is included in P. Thus P has at least one, and at most M members. To illustrate how the comparison between two distributions is performed, two normal distributions, X and Y, are considered.

Table I. Experimental results for various video sequences compared to HM 15.0

		Proposed Method			Gao et al. [2]			Park et al. [5]		
Class	*Video Sequences*	*TR(%)*	*BD-RATE(%)*	*BD-PSNR(dB)*	*TR(%)*	*BD-RATE(%)*	*BD-PSNR(dB)*	*TR(%)*	*BD-RATE(%)*	*BD-PSNR(dB)*
A	Traffic	-42.3	1.21	-0.057	-27.3	0.9	-0.05	-	-	-
	PeopleOnStreet	-42.7	1.44	-0.070	-24.7	0.9	-0.05	-	-	-
B	Cactus	-42.9	1.30	-0.043	-24.2	1.0	-0.03	-31.5	3.06	-0.11
	Kimono	-44.4	1.14	-0.037	-24.1	1.3	-0.04	-29.02	2.02	-0.06
	ParkScene	-41.6	0.79	-0.031	-26.1	0.7	-0.06	-25.45	2.38	-0.10
	BasketballDrive	-44.3	2.05	-0.052	-	-	-	-	-	-
	BQTerrace	-42.5	0.86	-0.042	-	-	-	-24.1	2.99	-0.17
C	BQMall	-42.1	1.21	-0.063	-34.2	0.9	-0.08	-29.01	2.06	-0.12
	PartyScene	-37.7	1.07	-0.072	-27.3	0.7	-0.05	-24.41	3.24	-0.25
	RaceHorsesC	-39.4	0.73	-0.041	-	-	-	-24.38	3.23	-0.20
	BasketballDrill	-42.1	0.77	-0.035	-28.6	1.2	-0.05	-30.57	3.24	-0.14
D	RaceHorses (Training)	-40.0	1.14	-0.065	-29.3	1.0	-0.06	-20.91	2.11	-0.14
	BasketballPass	-42.1	1.62	-0.086	-33.5	1.2	-0.08	-24.66	2.35	-0.13
	BlowingBubbles	-39.0	0.98	-0.051	-	-	-	-26.84	3.28	-0.20
	BQSquare	-38.6	1.48	-0.108	-26.7	1.1	-0.08	-28.32	2.39	-0.21
E	Vidyo1	-43.1	1.72	-0.075	-	-	-	-	-	-
	Vidyo3	-42.9	1.33	-0.064	-24.8	0.9	-0.04	-	-	-
	Vidyo4	-43.4	1.50	-0.060	-	-	-	-	-	-
Average		-41.8	1.24	-0.058	-27.6	1.0	-0.05	-26.6	2.70	-0.15

If X and Y are two independent normal distributions given by $X \sim N(\mu_x, \sigma_x^2)$ and $Y \sim N(\mu_y, \sigma_y^2)$, then

$$aX + bY \sim N(a\mu_x + b\mu_y, a^2\sigma_x^2 + b^2\sigma_y^2). \quad (5)$$

We are interested in computing $P(X < Y)$ while X and Y are related to m_{RMDmin} and a mode from set ψ, respectively, and compare it with a confidence level (CL). Essentially, we are evaluating the probability that a realization of X has a lower cost than a realization of Y. If this probability is lower than the desired confidence level, we add the mode associated with distribution Y to set P. Otherwise, we can safely exclude the mode associated with distribution Y from further processing and RDO cost computations. The confidence level provides a tradeoff between computational complexity and visual quality. To compute $P(X < Y)$ we define $W = X - Y$ as a new distribution, which based on (5), is normal with a mean $\mu_x - \mu_y$ and a variance $\sigma_x^2 + \sigma_y^2$ and consider $P(W < 0)$. Having the mean and variance of W, computing this probability becomes straightforward.

IV. Experimental Results

We implemented the proposed method in the HEVC test model HM15.0. The test platform was a PC equipped with an Intel® Core™ i7-4790 CPU @ 3.60 GHz and 32 GB of RAM. Since the proposed method was intended for intra coding, an *all-intra* configuration and *Main profile* were chosen. The experiments were conducted for the first 100 frames of seventeen 8-bit test sequences (non-training sequences) from five different video classes according to the definitions in [11]. Parameter K was set to 10, 10, 5, 5 and 5 for block sizes of 4×4, 8×8, 16×16, 32×32 and 64×64, respectively, and the confidence level was set to 0.65. An interesting feature of the proposed method is that increased speed or quality could be obtained by changing the confidence level, making it a very flexible method. To obtain the training data for RDO cost modeling, we used the *RaceHorses* sequence (class D), and as a result, the results for this sequence were not considered in the average results. To objectively compare the proposed method with the anchor HM 15.0, Bjontegaard delta bitrates (BD-RATE) and Bjontegaard delta peak signal-to-noise-ratios (BD-PSNR) were used [12]. For the computation of these criteria, we used four points of the rate-distortion curve. These points were achieved in our implementation by setting QP to 22, 27, 32 and 37.

Table I presents the time reduction (TR), BD-RATE and BD-PSNR for the proposed method and those in [2] and [5], as the most advanced and recently published works in this area, compared to HM 15.0. An empty cell in the table indicates that the authors did not report the results for the corresponding video sequence. It can be observed from this table that our proposed method achieves a 41.8% time reduction, with a 0.058 dB quality loss, while [2] and [5] provide time reductions of 27.6% and 26.6% with quality losses of 0.05 dB and 0.15 dB, respectively. Fig. 5 shows the rate-distortion curves of the proposed method and HM 15.0 for *BQMall*, indicating an insignificant loss of quality with our method. Similar curves are achieved for other sequences.

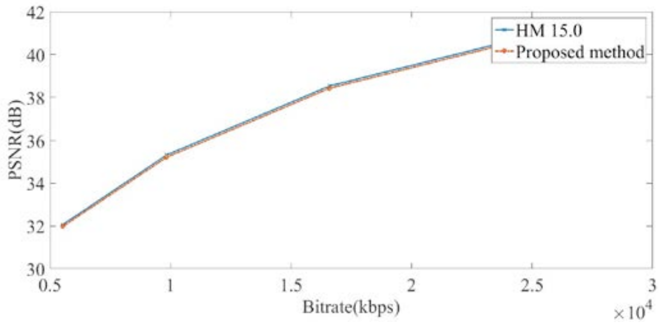

Fig. 5. RD curves of the proposed method and HM 15.0, *BQMall* sequence

V. Conclusion

In this paper, we have proposed a fast mode decision method based on RDO cost modeling to reduce the computational complexity of HEVC intra coding. The proposed method employs rough mode decision costs based on SATD to model the RDO costs using a normal distribution. This novel approach allows us to adaptively exclude from the RDO process the non-promising modes, i.e., those with a low chance of yielding the lowest RDO costs. Since the method relies on the actual RMD cost of candidates to decide which are non-promising, it can achieve high complexity reduction without sacrificing quality. Simulation results show that, on average, a 41.8% encoding time reduction is achieved using the proposed method, as compared to the reference HM 15.0, with a negligible BD-PSNR of 0.058 dB and 1.24% of BD-RATE.

References

[1] B. Bross, W. J. Han, J. R. Ohm, G. J. Sullivan, Y. K. Wang and T. Wiegand, "High Efficiency Video Coding (HEVC) Text Specification Draft 10", *Document JCTVC-L1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Jan. 2013.*

[2] L. Gao, S. Dong, W. Wang, R. Wang, W. Gao, "Fast Intra Mode Decision Algorithm Based on Refinement in HEVC", *in Circuits and Systems (ISCAS), IEEE International Symposium on, pp. 517-520, May 2015.*

[3] L. Wang and W. Siu, "Novel Adaptive Algorithm for Intra Prediction with Compromised Modes Skipping and Signaling Processes in HEVC", *in Circuits and Systems for Video Technology, IEEE Transactions on , vol. 23, no. 10, pp. 1686-1694, Oct. 2013.*

[4] G. Chen, Z. Pei, L. Sun, Z. Liu and T. Ikenaga, "Fast Intra Prediction for HEVC based on Pixel Gradient Statistics and Mode Refinement", *in Signal and Information Processing (ChinaSIP), IEEE China Summit & International Conference on, pp. 514-517, July 2013.*

[5] M. Park and J. Jeong, "Fast HEVC Intra Prediction Algorithm with Enhanced Intra Mode Grouping Based on Edge Detection in Transform Domain", *Journal of Advances in Computer Networks, vol. 3, no. 2, pp. 1686-1694, June 2015.*

[6] D. Gwon, Y. Choi and H. Choi, "Fast HEVC Intra Mode Decision Based on Bayesian Classification Framework with Relative SATD", *Springer-Verlag Berlin Heidelberg, vol. 330 of the series Lecture Notes in Electrical Engineering, pp. 203-209, 2015.*

[7] HEVC Test Model Software 15.0, https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/branches/

[8] V. Sze, M. Budagavi and G. J. Sullivan, "High Efficiency Video Coding (HEVC) Algorithms and Architectures", *Springer International Publishing Switzerland, 2014.*

[9] M. Jamali, S. Coulombe and F. Caron, "Fast HEVC Intra Mode Decision Based on Edge Detection and SATD Costs Classification", *in Proc. Data Compression Conference, pp. 43-52, Apr. 2015.*

[10] S. Ma, S. Wang, S. Wang, L. Zhao, Q. Yu, W. Gao, "Low Complexity Rate Distortion Optimization for HEVC", *in Proc. Data Compression Conference, pp. 73-82, Mar. 2013.*

[11] F. Bossen, "Common Test Conditions and Software Reference Configurations", *JCTVC-L1100, 12th Meeting, Geneva, Jan. 2013.*

[12] G. Bjontegaard, "Calculation of Average PSNR Differences between R-D Curves", *VCEG-M33, Tech. Rep, Apr. 2001.*