NEW TONE MAPPING AND TONE REPRODUCTION TECHNIQUES– APPLICATION TO BIT-DEPTH SCALABLE VIDEO CODING

Jui-Chiu Chiang, Che-Hsu Pan

Department of Electrical Engineering
National Chung Cheng University
Chia-Yi, Taiwan
Email: rachel@ccu.edu.tw, chpan@samlab.ee.ccu.edu.tw

ABSTRACT

With the advance of modern technology, we are looking for image/video representations with more realism. The popularity of HDTV (high definition TV) reveals that high resolution and high quality video soon becomes the milestone of the next video standard. Consequently, bit-depth scalability is likely becoming one important feature for the coming video coding standard. In this paper, we propose new tone mapping and tone reproduction techniques based on human visual perception. Then the proposed techniques will be applied on the bit-depth scalable video coding scheme. Experimental results show that the proposed coding scheme outperforms the prior work and the visual quality of low bit-depth images is promising.

1. GENERAL INFORMATION

High dynamic range (HDR) technologies for the capture and display of images/video content have matured rapidly in recent years. As a result, HDR imaging has become increasingly important in many applications, especially in the entertainment field, e.g., HDTV, digital cinema, mixed reality rendering, image/video editing, and remote sensing. In 2003, the Joint Video Team (JVT) called for proposals to enhance the bit-depth scope of H.264/AVC video coding [1]. The supported bit-depth in H.264/AVC is now up to 14 bits per colour channel. However, the bandwidth required to transmit encoded high bit-depth image/video content is much larger. Since not all display and printing devices support HDR images, rendering such images appropriately on conventional display devices is difficult. Tone mapping techniques have been developed to address the problem.

Tone mapping techniques can be classified into four categories [2], including global operations [3][4], local operations [5][6], gradient domain operations [7] and frequency-domain operations. Global methods produce low dynamic range (LDR) images according to some pre-defined tables or functions based on the feature of HDR images, but the methods also generate artefacts. The most significant artefacts result from distortion of the detail of the brightest or darkest area. Although such artefacts can be resolved by using a local operator, local methods are less popular than global methods due to their high complexity. In contrast, frequency domain operations emphasize compression of the low frequency content in an image; and gradient domain techniques try to attenuate the pixel intensity of areas with a high spatial gradient.

To cope with the increased size of high bit-depth image/video data compared to that of conventional LDR applications, it is necessary to develop appropriate compression techniques. Some approaches for HDR image compression that concentrate on backward compatibility with conventional image standards can be found in [8][9]. Moreover, to address the scalability issue, a number of bit-depth scalable video coding algorithms have been proposed in recent years, and many bit-depth related proposals have been submitted to JVT meetings [10-14]. Similar to spatial scalability in scalable video coding standard, the concept of inter-layer prediction is applied in bit-depth scalability to exploit the high correlation between bit-depth layers. For example, an inter-layer prediction scheme realized as an inverse tone mapping technique was proposed in [10]. The scheme predicts a high bit-depth pixel from the corresponding low bit-depth pixel through scaling plus offset, where the scale and offset values are estimated from spatial neighbouring blocks. Segall [15] introduced a bit-depth scalable video coding algorithm that is applied on the macroblock (MB) level. In this scheme, the base layer is also generated by tone mapping of the high bit-depth input and then encoded by H.264/AVC. For high bit-depth input, in addition to inter/intra prediction, inter-layer prediction is exploited to remove redundancy between bit-depth layers where a prediction from the low bit-depth layer is generated using a gain parameter and an offset parameter. Moreover, the high bit-depth layer and the low bit-depth layer use the same motion information estimated in the low bit-depth layer. In [11, 16], Winken *et al.* proposed a coding method that first converts a high bit-depth video sequence into a low bit-depth format, which is then encoded by H.264/AVC as the base layer. Next, the reconstructed base layer is processed inversely as a prediction mechanism to predict the high bit-depth layer. The difference between the original high bit-depth layer and the predicted layer is treated as an enhancement layer and no inter/intra prediction is performed for the high bit-depth layer. In our previous work [17], three bit-depth scalable video coding schemes were proposed. The proposed LH scheme is similar to most existing approaches because the high bit-depth layer is encoded by considering the inter-layer prediction of the corresponding low

 - ISSN 2076-1465

bit-depth layer, and its superiority over [15] was reported. Besides, the HL and LH-HL schemes in [17] differ from other prior works and the direction of inter-layer prediction is not always from the low-bit depth layer to the high bit-depth layer.

In this paper, we apply the proposed tone mapping and tone reproduction techniques on the LH scheme. The remainder of this paper is organized as follows. Section 2 reviews LH scheme, while Section 3 presents the proposed tone mapping and tone reproduction techniques. Section 4 details the experiment results. Then, in Section 5, we summarise our conclusions.

2. BIT-DEPTH SCALABLR CODING SCHEME

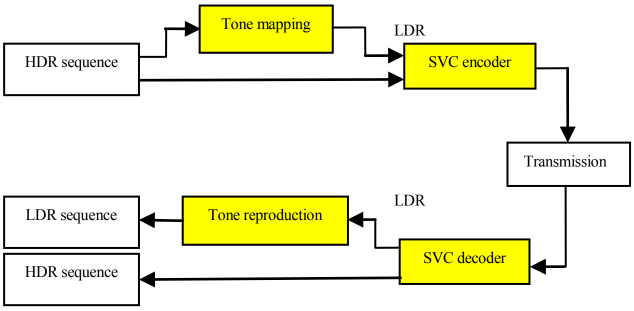


Figure1. The overall scheme

Figure 1 depicts the overall scheme. The HDR input is first tone mapped into a LDR format. Then both HDR and LDR sequence are encoded by the bit-depth scalable video encoder proposed in [17]. The HDR and LDR sequences are reconstructed by SVC decoder. Besides, the LDR sequence can be post-processed by tone reproduction technique to enhance the contrast and preserve the brightness.

Figure 2 shows the bit-depth scalable video coding scheme in [17]. Note that, "Reconstruction" module includes inverse quantization (IQ), inverse transformation (IT) and motion compensation. To ensure that the generated bitstream is embedded and compliant with the H.264/AVC standard, most bit-depth scalable coding schemes employ inter-layer prediction, which uses the low bit-depth layer to predict the high bit-depth layer [15-16]. The LH scheme in [17] adopts this idea with some modifications. We explain how it differs from other methods later in the paper.

The low bit-depth input is obtained after tone mapping of the original high bit-depth input and then encoded by H.264/AVC. Hence, the generated bit-depth scalable bitstream allows backward compatibility with H.264/AVC.

The right-hand side of Figure 2 shows the coding procedures for the high bit-depth layer. Like the low bit-depth layer, the encoding process is implemented on the MB level, but there are two differences. First, in addition to intra/inter prediction, the high bit-depth MB level gets another prediction from the corresponding low bit-depth MB by inverse tone mapping of the reconstructed low bit-depth MB. This prediction, which we call Intra Prediction from Low Bit-depth (IPLB), can be regarded as a type of inter-layer prediction and treated as an additional intra-prediction mode with a block size of 16×16, which is similar to inter-layer intra prediction performed in the spatial scalability of the SVC standard. Thus, two kinds of intra prediction are available in the proposed scheme: one explores the spatial redundancy within a frame, while the other tries to remove the redundancy between different bit-depth layers. Furthermore, to improve the coding efficiency of inter coding, the residual of the low bit-depth MB is inversely tone mapped and utilized to predict the residual of the high bit-depth MB. The process, called *Residual Prediction* can be regarded as another kind of inter-layer prediction and can be realized in two ways. The high bit-depth MB can perform motion estimation and motion compensation after subtracting the predicted residual from the original data. Such realization, called RP_1, is similar to inter-layer residual prediction in the spatial scalability of the SVC standard. Another realization of *Residual Prediction*, called RP_2, is to subtract the predicted residual after motion compensation. Both methods try to reduce the amount of redundancy in residuals of the low and the high bit-depth layers.

Basically, we utilize both IPLB prediction and *Residual Prediction* based on the results of R-D optimization. Note that there are four kinds of prediction in the proposed scheme: intra prediction, inter prediction, IPLB prediction, and *Residual Prediction*. Moreover, *Residual Prediction* cooperates with inter prediction if doing so yields better coding efficiency, while IPLB competes with other types of prediction. If inter-layer prediction (i.e., IPLB, RP_1 or RP_2) is not used, the high bit-depth layer is encoded by H.264/AVC. Next, we summarize the features of the LH scheme that distinguish it from several current approaches.

1. IPLB prediction: Similar to most bit-depth SVC schemes [15-16], the high bit-depth MB can be predicted from the corresponding low bit-depth MB by inverse tone mapping. However, in [16], intra/inter prediction is not realized in the high bit-depth layer, in conjunction with inter-layer prediction.
2. *Residual Prediction*: *Residual Prediction* can be applied in two ways. The high bit-depth MB can perform motion estimation after subtracting the predicted residual, or it can subtract the predicted residual after motion compensation. *Residual Prediction* is not used in the schemes proposed in [15][16]. The residual prediction operation described in [18] is performed only after motion compensation in the high bit-depth layer.
3. Motion Information: In the proposed scheme, both low bit-depth layer and the high bit-depth layer have their own motion information including the MB mode and motion vector (MV). This is different to the approach in [15], where the high bit-depth MB uses directly the motion information obtained in the corresponding low bit-depth MB

3. PROPOSED TONE MAPPING AND TONE REPRODUCTION TECHNIQUES

3.1 Proposed Tone Mapping Technique

In the proposed tone mapping technique, we first uniformly divide the whole intensity dynamic range (i.e., 0~2^{N-1}, where N is the bit-depth of the HDR image) into three parts (e.g.,

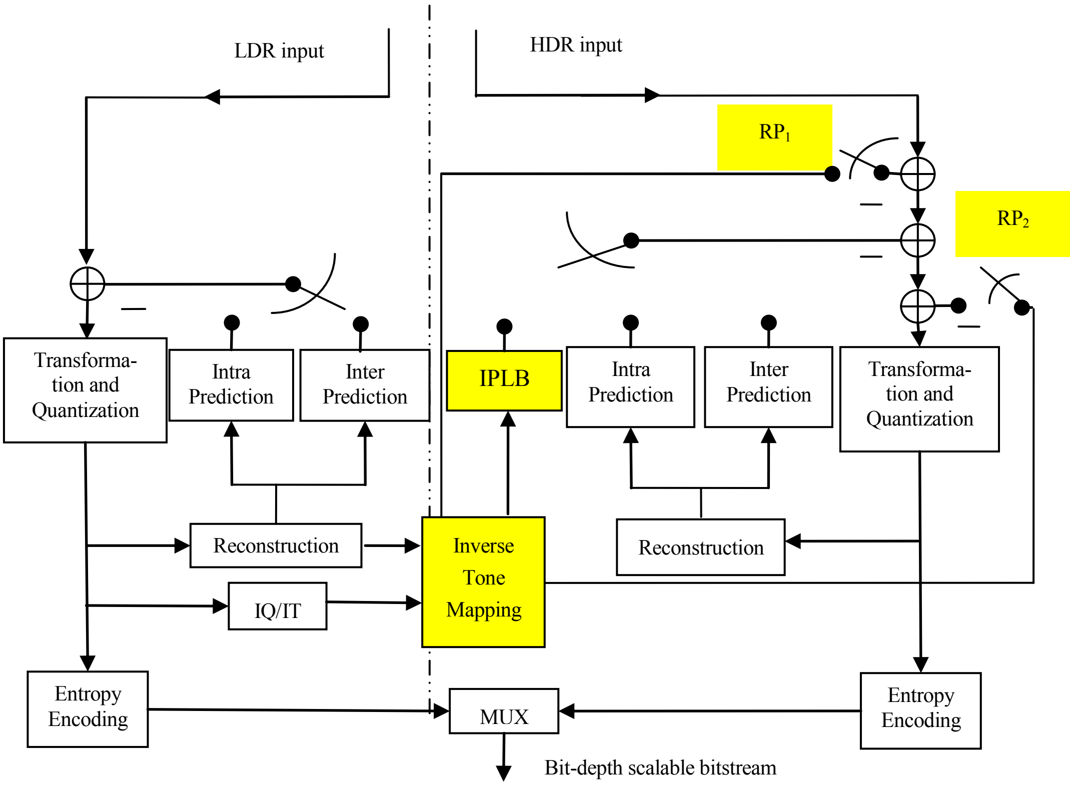


Figure 2. Bit-depth scalable coding scheme in [17].

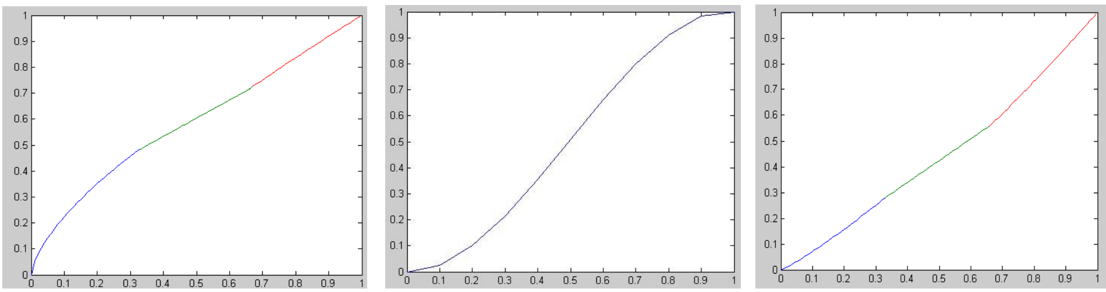


Figure 3. Different tone mapping curve

part_A, part_B, and part_C) and find the part with highest cumulative distribution function (CDF) and name it as part_Main.

Similar to the idea of histogram equalisation algorithm where more bins are assigned to the region with higher CDF, the proposed tone mapping technique will allocate more bins to part_Main. If part_Main is part_A, the proposed tone mapping formula operates as follows:

$$L = H^{\alpha} \quad \text{where} \begin{cases} \alpha = \alpha_1, \textit{for} \text{ part\_A} \\ \alpha = \alpha_2, \textit{for} \text{ part\_C} \end{cases} \tag{1}$$

where H is normalised HDR pixel intensity, L is the normalised LDR pixel intensity. In this case, the pixel intensity in part_B is obtained by linear mapping from $(1/3)^{\alpha 1}$ to $(1/3)^{\alpha 2}$. Figure 3(a) presents the tone mapping curve of this case. It indicates that more intensity levels are assigned to part_A.

Due to the physiology observation that the sensitivity to luminance intensity for human eyes is *S*-shape like, the tone mapping formula when part_Main is part_B follows this rule and can be expressed as,

$$L = -2.2024 \times H^3 + 3.2679 \times H^2 + 0.0655 \times H \tag{2}$$

Figure 3(b) presents the tone mapping curve of this case. This curve is *S*-shape like. However, the shape is not very sharp to avoid over compression in both high and low luminance parts.

If part_Main is part_C, the proposed tone mapping formula performs as follows:

$$L = H^{\beta} \quad \text{where} \begin{cases} \beta = \beta_1, \textit{for} \text{ part\_A} \\ \beta = \beta_2, \textit{for} \text{ part\_C} \end{cases} \tag{3}$$

Figure 4. Tone mapping results (left: [3], middle: [5], right: proposed)

In this case, the pixel intensity in part_B is obtained by linear mapping from $(1/3)^{\beta 1}$ to $(1/3)^{\beta 2}$. Figure 3(c) presents the tone mapping curve of this case. It indicates more intensity levels are assigned to part_C.

3.2 Proposed Tone Reproduction Technique

The proposed tone reproduction algorithm will enhance the contrast of the LDR image by referring the contrast in the HDR image. First, the contrast between two pixels *m*, *n* is defined as,

$$C(m,n)=\left|\frac{m-n}{m+n}\right| \tag{4}$$

We then modify the LDR pixel value by minimising the contrast difference between the LDR content and the corresponding HDR content as follows [19]:

$$y' = \arg min\left|\sum_{i=1}^{4} C(y,y_i) - \sum_{i=1}^{4} C(x,x_i)\right| \tag{5}$$

where *y, x* denotes respectively, the original LDR and the corresponding HDR pixel values and y_i, x_i , i={1,2,3,4} denote their surrounding pixel values. Besides, the refinement window for y' is ± 8 around the value of *y*. Then the final LDR value y_{final} is obtained as shown in (6) by considering both the original LDR value and y' to make the intensity more natural.

$$y_{\text{final}} = (2y + y')/3 \tag{6}$$

4. EXPERIMENTAL RESULTS

In this section, we present experimental results of our proposed methods. The parameters used are: α_1=0.6, α_2=0.9, β_1=1.1 and β_2=1.4. The proposed tone mapping and tone reproduction techniques are applied on the luminance component, while the chrominance components are tone mapped by bit truncation. Figure 4 shows the tone mapping results of the proposed technique, as well as the method proposed in [3] and [5]. It reveals that the proposed method preserves global contrast and presents good visual quality.

To evaluate the coding efficiency brought by the proposed tone mapping technique, the bit-depth scalable coding scheme in [17] is implemented on the SVC reference software JSVM (9.16). The HDR input is 12-bit video sequence Sunrise (960×540) and Library (900×540) at 30 Hz [20], with 420 format while the LDR is 8-bit. The two bit-depth layers use the same quantization parameter (QP), and the inverse tone mapping is carried out by table looking-up where the mean squared error between the original 12-bit data and the reconstructed 12-bit data after tone mapping and inverse tone mapping is minimised. The QPs used here are {24, 28, 32, 36, and 40}.

Figure 5 shows the coding performance for sequence “Sunrise”. Table 1 lists the comparison in terms of Bjontegaard metric [21] for the high bit-depth layer in both sequences. Figure 5 and Table 1 show that the bit-depth scalable coding proposed with the proposed tone mapping technique has better coding efficiency in both bit-depth layers. In particular, the bitrate saving in the LDR layer is significant. It implies that the proposed tone mapping technique not only makes the visual quality satisfied, as indicated in Figure 4, but also provides an improved coding efficiency than [5] in video compression.

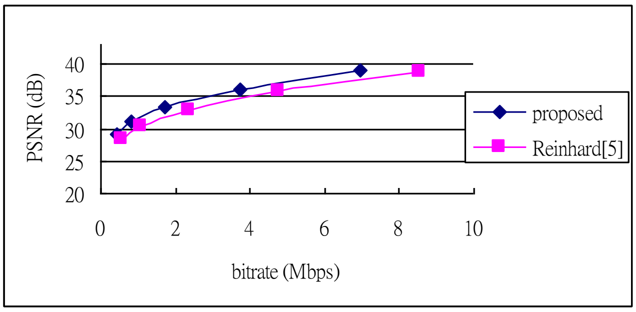
The performance of the proposed tone reproduction technique is illustrated in Figure 6. It reveals more details outside the windows are enhanced after the processing of tone reproduction.

5. CONCLUSION

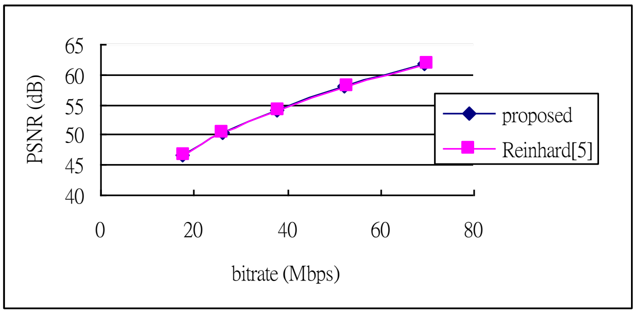
In this paper, we proposed a new tone mapping technique to well preserve the characteristics of the HDR image on the tone mapped LDR image. The experimental result shows the visual quality of the LDR image generated by the proposed technique is quite good. Moreover, the proposed tone mapping technique is also performed on the bit-depth scalable video coding scheme. Experimental results show that the improved coding efficiency brought by the proposed tone mapping technique is obvious. Besides, to further enhance the visual quality of the LDR image, a new tone reproduction technique was proposed. The effectiveness of the proposed tone reproduction technique is also demonstrated by exhibiting more details on the processed LDR image.

Table 1. BDBR& BDPSNR of the proposed technique with respect to [5] for "Sunrise" (GOP=8).

	BDBR (%)	BDPSNR(dB)
Sunrise (12-bit)	-1.03	0.12
Library (12-bit)	-2.72	0.24



(a) 8-bit "Sunrise" sequence. (GOP=8)



(b) 12-bit "Sunrise" sequence. (GOP=8)

Figure 5. R-D performance of the proposed technique.

(a) before tone reproduction

(b) after tone reproduction

Figure 6. Image illustrations for the proposed tone reproduction.

REFERENCES

[1] G. Sullivan, A. Luthra, and T. Wiegand, "Call for Proposals for Extended Sample Bit Depth and Chroma Format Support in the Advanced Video Coding Standard," *Joint Video Team*, Doc. JVT-G048, Pattaya II, Thailand, March 2003.

[2] E. Reinhard, S. Pattanaik, G. Ward, and P. Debevec, "High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting," Morgan Kaufmann Publisher, 2006.

[3] J. Duan and G. Qiu, "Fast Tone Mapping For High Dynamic Range Images", *in Proc. of 17th Int. Conf. on Pattern Recognition*, vol. 2, 2004. pp. 847-850.

[4] F. Drago, K. Myszkowski, T. Annen and N. Chiba, "Adaptive Logarithmic Mapping For Displaying High Contrast Scenes", *The Journal of Computer Graphics Forum*, vol.22, no. 3, pp. 419-426, 2003.

[5] E. Reinhard, M. Stark, P. Shirley and J .Ferwerda, "Photographic Tone Reproduction for Digital Images," in *Proc. of ACM SIGGRAPH,* 2002, pp. 267-276.

[6] F. Durand and J. Dorsey, "Fast Bilateral Filtering for The Display of High-Dynamic-Range images," in *Proc. of ACM Trans. Graph.* (*Special issue SIGGRAPH 2002*) vol. 21, no. 3, pp. 257-266, 2002.

[7] R. Fattal, D. Lischinski and M. Werman, "Gradient Domain High Dynamic Range Compression," in *Proc. of ACM SIGGRAPH 2002*, 2002. pp. 249-256.

[8] G. Ward, and M. Simmons, "JPEG-HDR: a Backward-compatible, High Dynamic Range Extension to JPEG," in *Proc. of the 13th Color Imaging Conference*, Nov. 2005.

[9] M. Okuda, and N. Adami, "Two-layer Coding Algorithm for High Dynamic Range Images Based on Luminance Compensation," *Journal of Visual Communication & Image Representation*, vol. 17, pp. 377-386, 2007.

[10] S. Liu, A.Vetro and W. S. Kim, "Inter-layer prediction for SVC Bit-depth Scalability," *Joint Video Team*, Doc. JVT-X075, Geneva, Switzerland, June 2007.

[11] M. Winken, H. Schwarz, D. Marpe, and T. Wiegand, "SVC Bit Depth Scalability," *Joint Video Team,* Doc. JVT-V078. Marrakech, Morocco, Jan. 2007.

[12] A. Segall, and Y. Su, "System for Bit-Depth Scalable Coding," *Joint Video Team,* Doc. JVT-W113. San Jose, California, USA, April 2007.

[13] Y. Ye, H. Chung, M. Karczewicz, and I. S. Chong, "Improvement to Bit Depth Scalability Coding," *Joint Video Team*, Doc. JVT-Y048, Shenzhen, China, Oct, 2007.

[14] Y. Yu, S. Gordon, and M. Yang, "Improving Compression Performance in Bit Depth SVC with a Prediction Filter," *Joint Video Team,* Doc. JVT-Z045, Antalya, Turkey, Jan. 2008.

[15] A. Segall, "Scalable Coding of High Dynamic Range Video," in *Proc. of IEEE International Conference on Image Processing,* 2007, pp.1-4.

[16] M. Winken, D. Marpe, H. Schwarz, and T. Wiegand, "Bit-depth Scalable Video Coding," in *Proc. of IEEE International Conference on Image Processing,* 2007, pp.5-8.

[17] J. -C. Chiang, and W. -T. Kuo, "Bit-depth Scalable Video Coding using Inter-layer Prediction from High Bit-depth Layer," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing,* 2009, pp. 649-652.

[18] Y. Wu, Y. Gao, and Y. Chen, "Bit-depth Scalable Coding Based on Macroblock Level Inter-layer Prediction," in *Proc. of IEEE Symposium Conference on Circuits and Systems*, 2008, pp.3442-3445.

[19] K.-Y. Yip, O. C. Au, "Optimal Contrast Enhancement for Tone-mapped Low Dynamic Range Images Based on High Dynamic Range Images", in *Proc. of PACRIM'09*, Communications, Computers and Signal Processing, 2009, pp. 53-58.

[20] A. Segall "Donation of Tone Mapped Image Sequences," *Joint Video Team*, Doc. JVT-Y072, Shenzhen, China, October, 2007.

[21] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-curves," *13th VCEG-M33 Meeting*, Austin, TX, Apr., 2001.