

Asymmetric Stereoscopic Images Coding using Perceptual Model

Samira Ouddane, Kamel Mohamed Faraoun
Djillali Liabes University of Sidi Bel Abbès, Department of Computer Science
Sidi Bel Abbès, Algeria
ouddane.s@gmail.com



ABSTRACT: Significant research works have been devoted to asymmetric stereoscopic coding technique. Based on the perceptual theory of binocular suppression, in this method, one view of the stereo pair can be coded with a lower quality than the other, thus allowing to decrease bandwidth requirement for 3D visual content delivery without degrading subjective quality. However, with the condition that the quality gap between both images does not exceed a certain threshold. Some works proposed to fix this threshold by subjective experiments. Nevertheless, using a fixed threshold does not allow adaptation to the stereo image content and presents dependency to the experimental design. In this paper, we propose a novel method that selects automatically the maximum tolerable quality gap between views for different stereo image pairs. This has been achieved by modeling the relationship between the asymmetric quality and the interview distortion between corresponding pixels. Then, the derived relationship is combined with Binocular Just Noticeable Difference (BJND) model to control the quality threshold between the left and right views. Experimental results showed that the proposed coding method can achieve considerable bitrate saving up to 16.32% and outperforms the widely used asymmetric coding approaches.

Keywords: Stereoscopic Images, Asymmetric Coding, Binocular Suppression, Binocular Just-noticeable Difference, Human Visual System

Received: 12 December 2017, Revised 19 January 2018, Accepted 5 February 2018

DOI: 10.6025/jmpt/2018/9/2/59-70

© 2018 DLINE. All Rights Reserved

1. Introduction

In recent years, the provided digital multimedia content is very rich and the multimedia systems have evolved exponentially. These aimed to increase the quality of experience (QoE) of the users by offering them a realistic feeling of immersion. The most obvious example is the 3D visual media applications that provide the illusion of depth perception of the scene.

Most of the 3D content is captured and prepared in stereoscopic format [1]. Stereoscopic 3D (S-3D) consists of capturing two images from two slightly different positions and provides each of them to each viewer's eye. Thus, stereoscopic content will require twice the amount of data compared to monoscopic one. In view of the huge amount of data involved by media transmission and consumption of such S-3D content, the development of efficient compression techniques is of paramount importance.

The easy and obvious way to stereoscopic images coding is to encode each view of the stereo pair independently, using any still image compression standard, such as JPEG [2] or JPEG2000 [3]. This approach is usually referred as simulcast coding method and offers some advantages such as low complexity, an easy deployment and backward compatibility with 2D systems. However, since the redundancy between views (i.e., interview redundancy) is not exploited in this approach, the coding performance is clearly suboptimal. Indeed, the views of a stereo pair represent the same scene captured from different viewpoints and consequently contain a large amount of interview statistical dependencies. Therefore, stereoscopic images coding can be performed more efficiently by exploiting such interview correlations [4, 21]. This is usually achieved by predicting one view from the other through the so-called disparity estimation and compensation mechanism. More specifically, one view of the stereo pair is selected as the reference image, and the other view is considered as the target one. Thereafter, the disparity map is calculated to be used in predicting the target image from the reference one. Next, the difference between the predicted image and the original target image generates the so-called residual image, which is transformed, quantized, encoded and finally transmitted with the disparity map.

A considerable research efforts have been dedicated to enhance the different steps mentioned above [5]. Some of them tackled the transformation step when coding the reference and residual images [6–10], for instance, some works showed that the residual image has specific characteristics to be exploited by transformation step [6,7]. Other contributions proposed to apply different types of transformation to the occluded and non-occluded blocks [10]. On the other hand, other works aimed at enhancing the disparity estimation and compensation [11–14]. To estimate the disparity between the reference and target images, in [11], the authors proposed a bandelet-based approach using geometrical properties of both images. In [12], an overlapped block-based matching method was proposed, including smoothness constraints and using adaptive windows with variable shapes.

Since the human observer is the ultimate receiver, it would be advantageous to integrate the human visual system (HVS) properties to enhance the coding performance. Following this way, some perceptual approaches exploiting the perceptual properties of the HVS have been proposed for stereoscopic images coding. Among them, we can cite the asymmetric coding techniques that are based on the so-called binocular suppression theory [15]. The latter specifies that in the case where the eyes are provided with two views having a different quality, the fused 3D image will be dominated by the higher quality view [16]. Based on this theory, in order to reduce the required bitrate for stereoscopic images delivery, one image of the stereo pair can be encoded at high quality level, while the other image is encoded at a slightly lower quality, without noticeable visual quality losses.

Many asymmetric coding methods have been proposed in the literature [15]– [27], the main difference between the different methods is the means by which the asymmetry in quality between the two views is achieved. For instance, some methods proposed to encode both views of the stereo pair at unequal quality or bitrates, i.e., at unequal quantization parameter (QP); this method is referred to as quality ($PSNR$) reduction or asymmetric quantization [16]– [19]. Implementing asymmetry by quality/ $PSNR$ reduction is straightforward and keeps the encoder/decoder unchanged. This is done by coding one view of the stereo pair more heavily than the other, thus one of views shows more compression artifacts. However, the fused perceptual 3D image is dominated by the higher quality view. For instance, in order to decrease the overall bitrate, Saygili et al. [18] proposed to encode one view at sufficiently high quality ($PSNR = 40$ dB), while the quality of the secondary view is reduced but kept above a certain $PSNR$ threshold. They observed that the degradation in 3D perception is unnoticeable and the perceptual quality of 3D image is driven by the high quality view. This can be explained by the fact that the high quality image masks the effect of the lower quality one.

With the same aim of reducing the transmission bandwidth, other asymmetric coding methods proposed to present both views with different spatial resolutions. This approach is usually known as the mixed resolution (MR) method. Therefore, in the MR concept, before the encoding stage, one of the views is low-pass filtered (blurred image), while the other view is kept at its original full resolution (sharp image). Due to the smaller number of pixels to be encoded compared to the full resolution case, this operation leads to reduce the transmission bandwidth. At the decoding stage, the view with lowest resolution is up-sampled to the full resolution. Finally, the perceived 3D fusion of the two images (sharp and blurred images) is relatively close to the sharper image [20]– [27].

The asymmetric stereoscopic image coding has demonstrated its effectiveness to provide significant bitrate saving without noticeable degradation, up to a certain limit. From another perspective, some works tried to quantify this perceptual limit of

asymmetry in the context of stereoscopic image/video coding, and they used the obtained limit of asymmetry to perform effective asymmetric coding in such a way that the viewer cannot perceive any degradation. For instance, in [18], the authors found that the tolerable difference between both views in term of PSNR, is about 9 dB for a parallax barrier display and 7 dB for a full resolution projection display. In the work described in [27], 2 dB of difference between the left and right views has been considered as a maximum tolerance level for which the viewer does not perceive any annoying distortions. Finally, for asymmetric coding by reduction of spatial resolution, the range of downsampling ratios between 1/2 and 3/8 has showed satisfactory 3D viewing experience [22].

Despite these valuable contributions, all of these latter studies described above are based on subjective experiments and depend on the experiment design, consequently, making their generalization difficult. In addition, affecting a fixed value to the limit of asymmetry, does not allow any adaptation to the content.

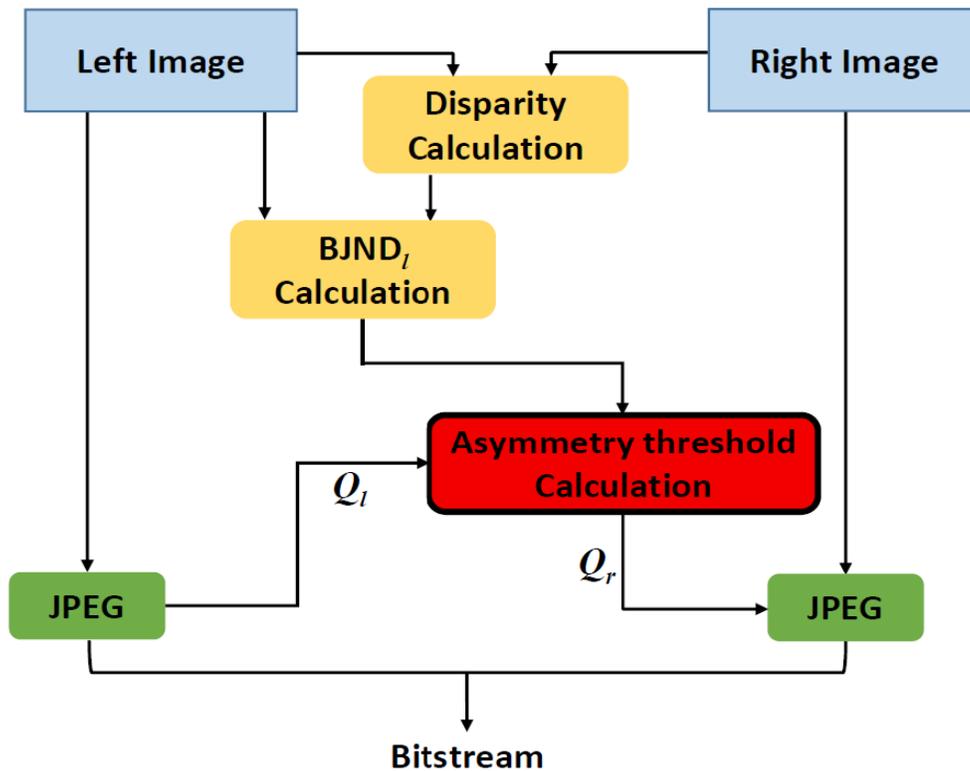


Figure 1. The proposed framework for asymmetric stereoscopic images coding method

Consequently, in this paper, we propose a novel method that determines automatically the most suitable gap in quality between the left and the right images in asymmetric stereoscopic images coding, which provides the best 3D viewing experience. To achieve this perceptual distortion control, the relationship between the asymmetric quality and the interview distortion between corresponding pixels is expressed as a function relying on the binocular just-noticeable difference (*BJND*) [29] model.

The rest of this paper is organized as follows. Section 2 describes the proposed approach as well as the exploited *BJND* model. Experimental results are presented in Section 3. Finally, Section 4 concludes this paper and provides direction for the future works.

2. Proposed Approach

As described previously, when the left and right images are presented to the viewer with different qualities, thanks to the binocular suppression phenomenon of the HVS, the perceptual quality in 3D viewing is dominated by the high quality view. Therefore, based on this concept, we can compress one view of the stereo pair more heavily (with a lower bit-budget) than the

other view without introducing visible artifacts during 3D viewing, provided that the quality difference between both views does not exceed a given threshold.

In the following, we describe the proposed asymmetric stereoscopic images coding that adjusts automatically and adaptively the quality threshold between both views according the perceptual model, making the quality difference almost transparent to the viewer, thus offering optimal 3D visual experience.

The framework of the proposed method is shown in Figure 1. Given a stereo pair composed of the left I_l and right I_r images. In the rest of this paper, we consider I_l as the reference image and I_r as the target one. In order to achieve bitrate saving, I_r is the one that we strongly compress and thus has the lowest quality.

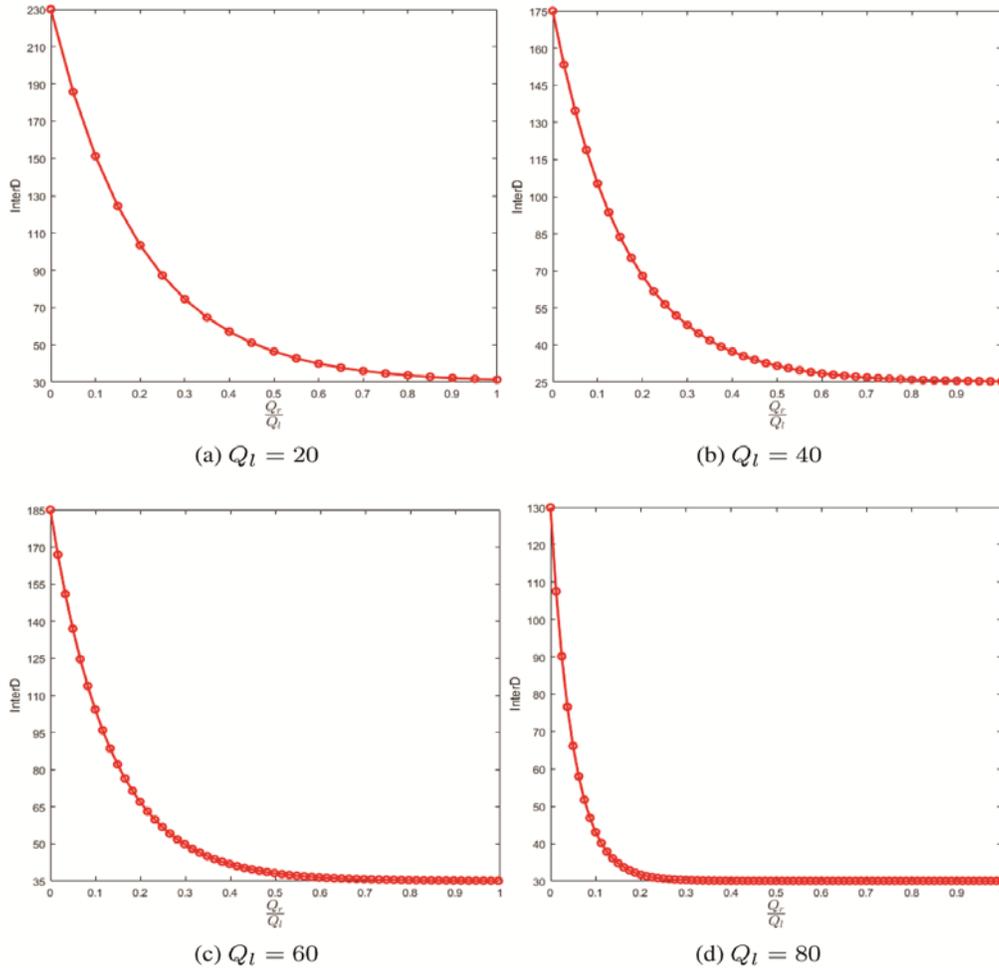


Figure 2. The curve distributions of $InterD$ and Q at different Q_l and Q_r values, where Q_r and $InterD$ are computed by Eq.(3) and Eq.(2), respectively

To be able to adjust the asymmetric quality of the stereo pair, i.e., tuning the quality of $I_r (Q_r)$ relative to that of $I_l (Q_l)$, it is important to model the relationship between the quality of both views (I) and their interview distortion ($InterD$). It is assumed that the $InterD-Q$ relationship can be described by a function as follows:

$$InterD = f(Q) \tag{1}$$

where Q represents the ratio between the qualities of both views (Q_r / Q_l), ranging from [0; 1]. While $InterD$ represents the mean-square-error distortion measure between corresponding pixels of both views, which is defined as

$$InterD = \frac{1}{n} \sum_{i=1}^n (I_l(x, y) - I_r(x, y + d_i))^2 \quad (2)$$

where n is the number of matched pixels between I_l and I_r . $I_l(x, y)$ denotes the luminance value of the pixel i at position (x, y) . d_i is the disparity value of the pixel i .

Next, in order to define the function $f(\cdot)$, the $InterD - Q$ relationship at different Q_r / Q_l values is evaluated by experiment way using a set of different stereoscopic images [30]. For each stereo pair, we used ten different Q_l values ranging from [10, 100] with a step of 10, and varied the Q_r values on the basis of Q_l as follows

$$Q_r = Q_l - i, \text{ with } i \in [0, Q_l] \quad (3)$$

Figure 2 shows the curve distributions of $InterD$ and Q at different Q_l and Q_r values for Aloe stereo pair. Based on the obtained results, we can observe that the general shape of the distribution follows an exponential curve. Consequently, the function describing the $InterD - Q$ relationship in Eq.(1) can be formulated as

$$InterD = \alpha_1 \exp\left(-\alpha_2 \cdot \frac{Q_r}{Q_l}\right) \quad (4)$$

where α_1 and α_2 are the model parameters determined experimentally. As mentioned previously, the aim is to decrease the quality in one view without introducing visible artifacts in 3D viewing. In other words, decreasing Q_r should lead to a change below the binocular visibility threshold ($BJND$) [29]. In the following, we give a brief overview of the $BJND$ model.

The Just Noticeable Difference (JND) can be defined as the minimum change that could be noticed by a standard viewer. In other words, JND refers to the visibility threshold below which any change cannot be detected by the HVS [28]. In order to build a Binocular- JND ($BJND$) model, Zhao et al. conducted psychophysical experiments [29]. Based on these experiments, they proposed a $BJND$ model to measure the joint visibility of a pair of distortions in two views. The $BJND$ model determines the minimum distortions in one view that evoke binocularly visible differences, given the background information and the distortions in the corresponding area of the other view.

In their proposed model, they considered two HVS characteristics namely luminance and contrast masking effects, and they modeled them to the case of binocular vision as described in the following.

Given the left and right views, as well as the disparity map of the left image. The $BJND$ of the left view ($BJND_l$) is defined as:

$$\begin{aligned} (BJND_l)(i, j) &= (BJND_l)(bg_r(i - d_p j), eh_r(i - d_p j), n_r(i - d_p j)) \\ &= A_c(bg_r(i - d_p j), eh_r(i - d_p j)) \\ &\quad \times \left(1 - \left(\frac{n_r(i - d_p j)}{A_c(bg_r(i - d_p j), eh_r(i - d_p j))}\right)^\lambda\right)^{1/\lambda} \end{aligned} \quad (5)$$

where i and j are the pixel coordinates, d_i is the horizontal disparity value at pixel (i, j) . The parameter λ controls the influence of the noise in the right image, and it was suggested in [29] that $\lambda = 1.25$. It should be noted that $BJND_l$ is dependent on the background luminance bg_r , the edge height eh_r , and the noise amplitude n_r of the corresponding pixel in the right image. Where no noise in the right image, i.e., $n_r(i - d_p j) = 0$, the $BJND_l$ is reduced to A_c , which is defined as

$$A_c(bg, eh) = A_{limit}(bg) + K(bg) \cdot eh \quad (6)$$

Thanks to psychophysical experiments, authors defined $A_{limit}(bg)$ and $K(bg)$, respectively by

$$A_{limit}(bg) = \begin{cases} 0.00027 \cdot (bg^2 - 96 \cdot bg) + 8, & \text{if } 0 \leq bg < 48 \\ 0.0001 \cdot (bg^2 - 32 \cdot bg) + 1.7, & \text{if } 48 \leq bg \leq 255 \end{cases} \quad (7)$$

$$K(bg) = -10^{-6} \cdot (0.7 \cdot bg^2 + 32 \cdot bg) + 0.07 \quad (8)$$

where bg is the mean of the luminance values of a block of 5×5 centered on the corresponding pixel position, and the edge height eh is computed by the 5×5 Sobel operators as follows:

$$eh(i, j) = \sqrt{E_H^2(i, j) + E_V^2(i, j)} \quad (9)$$

$$E_k(i, j) = \frac{1}{24} \sum_{h=1}^5 \sum_{v=1}^5 I(i-3+h, j-3+v) \cdot G_k(h, v), \quad (10)$$

with $k = H, V$.

where the detailed representation of the Sobel operator $G_k(h, v)$ is as follows

$$G_H = \begin{pmatrix} -1 & -2 & 0 & 2 & 1 \\ -2 & -3 & 0 & 3 & 2 \\ -3 & -5 & 0 & 5 & 3 \\ -2 & -3 & 0 & 3 & 2 \\ -1 & -2 & 0 & 2 & 1 \end{pmatrix}, \quad G_V = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 2 & 3 & 5 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -3 & -5 & -3 & -2 \\ -1 & -2 & -3 & -2 & -1 \end{pmatrix} \quad (11)$$

To the best of our knowledge, this is the first model to measure the perceptible distortion threshold of stereoscopic images for binocular vision, and we exploited it in the proposed method to ensure reliable 3D QoE.

Thus, the perceptual distortion control is achieved in the proposed method, by restricting the interview distortion (i.e., $InterD$) between views by $BJND$ as follows

$$InterD < BJND \quad (12)$$

Therefore, if we substitute $InterD$ in Eq.(12) by its definition in Eq.(4), we can derive automatically the appropriate value for Q_r as follows

$$Q_r > -\frac{Q_l}{\alpha_2} \cdot \ln\left(\frac{BJND}{\alpha_1}\right) \quad (13)$$

In this way, to avoid any visible distortion when exploiting the asymmetric coding method, the selected value of Q_r should not be below the threshold defined in the Eq.(13). In other words, according to the Q_l value in addition to the $BJND$ which is calculated relative to the content of both views, we can automatically derive the optimal Q_r value that achieves significant bit-rate saving while providing optimal quality of experience.

3. Experimental Results

In this section, simulation results are provided to evaluate the performance of the proposed asymmetric stereoscopic images coding method. Evaluations have been carried out on the Middlebury stereo dataset [30]. We used eight stereo image pairs with different resolutions ranging from 450375 to 13901110, named Art, Books, Cloth3, Cones, Midd1, Midd2, Moebius and Teddy. Figure 3 shows the left image of each stereo pair considered in the performance evaluation.

The proposed method is dependent on α_1 and α_2 parameters (see Eq.(13)). The latter are determined through a deep experimentation

and, accordingly, the optimal results have been obtained when setting α_1 to 25 and α_2 to 0.35, given that $BJND$ values do not exceed 20.

The propose method was compared with symmetric stereoscopic images coding method, which is considered as a reference or anchor coding method (denoted as SC). For the latter, both views have been coded in the same way, without favoring one view to the detriment of the other. In addition, three widely used asymmetric stereoscopic images coding methods have also been considered in the comparison. The first one achieves the asymmetry by downsampling one view with a ratio of 1/2 along both coordinate axes (denoted as AC_Down), and the second consists in blurring one view of the stereo pair using Gaussian blur filter (denoted as AC_Blur) and, finally, in the third method, both views are coded at unequal quality, with a fixed quality gap of 10 between them (denoted as AC_FQuality).

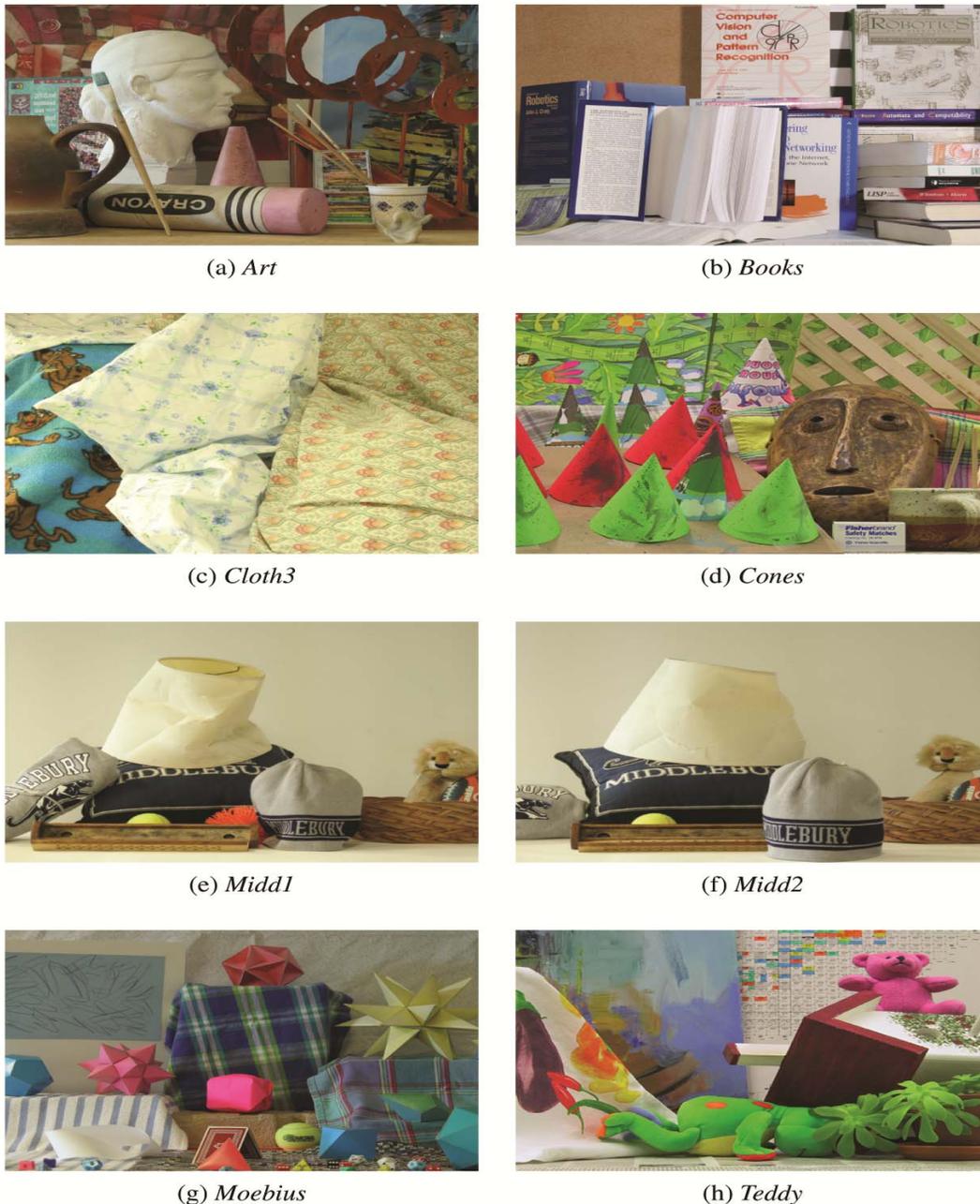


Figure 3. The eight stereoscopic images used in the experiment. Shown here are only the left images

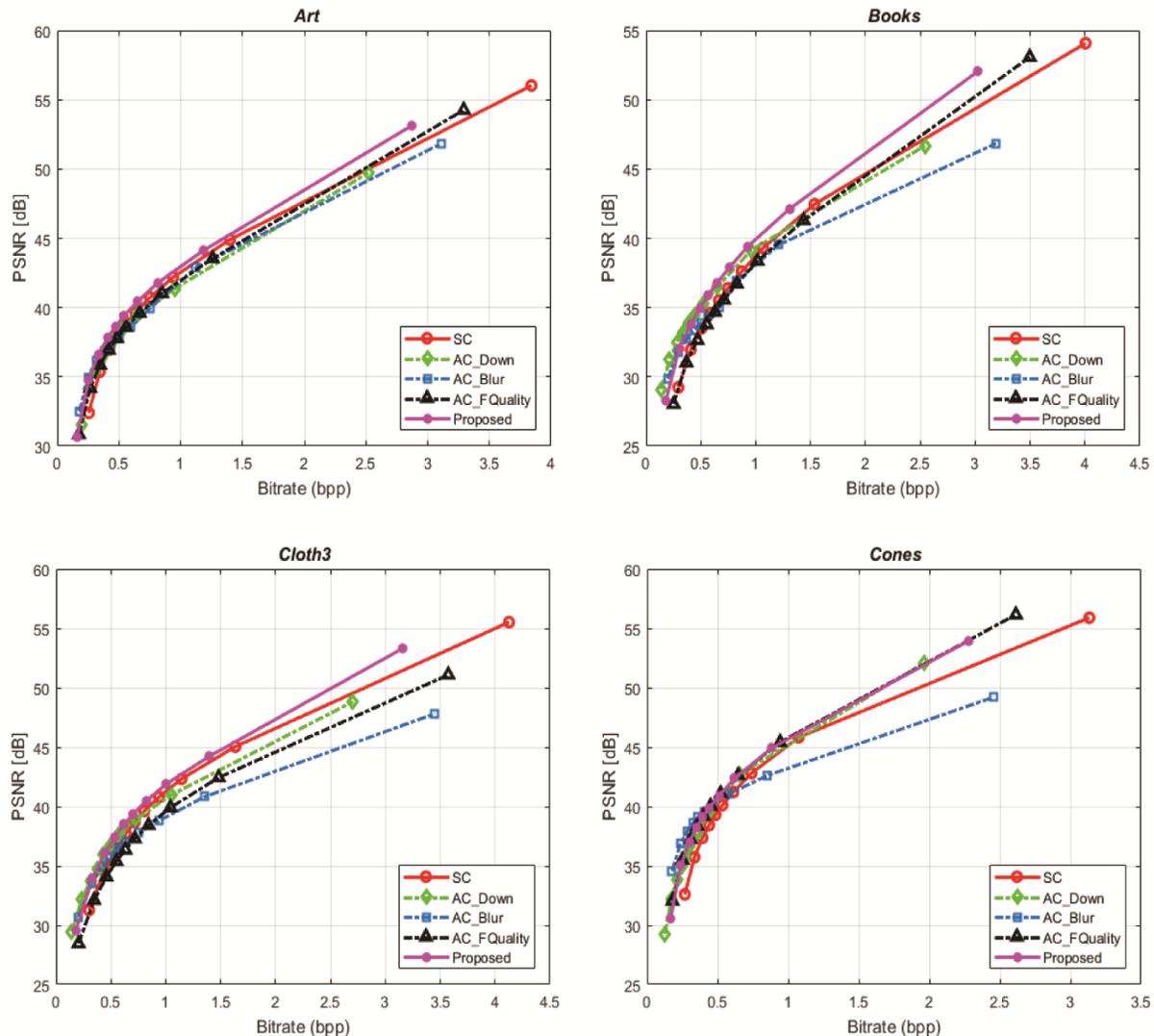
For all compared methods, we used *JPEG* [2] as the coding standard, with ten different quality values ranging from [10, 100]. As mentioned previously, the left image is coded as the reference and the right one as the target. The results are presented as a quantitative comparison in terms of average bitrate and its corresponding *PSNR* measure. The latter is computed as an average of the mean squared error (*MSE*) of the left and right reconstructed images as follows:

$$PSNR = 10 \cdot \log_{10} \left(\frac{255^2}{(MSE_l + MSE_r)/2} \right)$$

$$\text{with } MSE_k = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M (I_k(i,j) - \hat{I}_k(i,j))^2, k \in \{l, r\}$$
(14)

where (i, j) are the pixel coordinates and (N, M) are the dimensions of the images. I and \hat{I} denote the original and reconstructed (decoded version) images, respectively.

The coding performance are illustrated through the rate-distortion (*R-D*) curves for different images, as shown in Figure 4. The vertical axis represents the average *PSNR* and the horizontal axis reports the bitrate required to encode the stereoscopic images. In addition, the average bitrate savings, which are calculated according to the widely used Bjøntegaard measurement [31], are reported in Table 1.



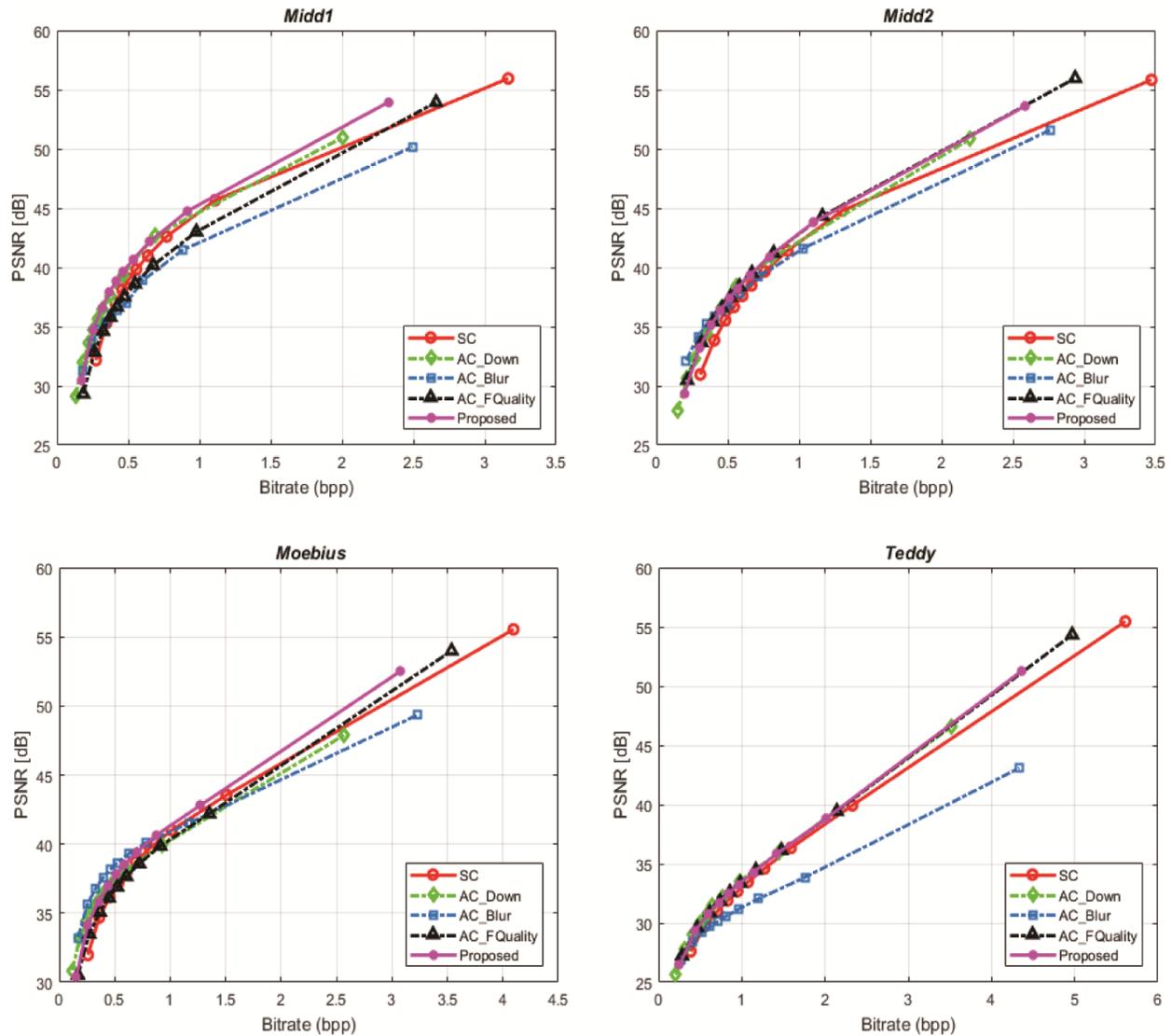


Figure 4. Rate/distortion curves for the eight used stereoscopic image pairs

| Images | AC_Down | | AC_Blur | | AC_FQuality | | Proposed | |
|----------------|----------------|---------------|----------------|---------------|----------------|---------------|----------------|---------------|
| | Bitrate saving | Δ PSNR |
| <i>Art</i> | -13.06% | 0.56 | -8.40% | 0.61 | -6.16% | 0.74 | -14.66% | 1.07 |
| <i>Books</i> | -11.20% | 0.94 | -3.32% | 0.05 | -2.66% | 0.12 | -15.32% | 1.23 |
| <i>Cloth3</i> | -5.64% | 0.84 | 1.85% | 0.03 | 1.20% | -0.77 | -8.88% | 0.98 |
| <i>Cones</i> | -15.31% | 1.09 | -12.42% | 0.30 | -12.49% | 0.89 | -15.90% | 1.55 |
| <i>Midd1</i> | -13.35% | 1.03 | 0.97% | 0.14 | -6.61% | 0.20 | -15.25% | 1.13 |
| <i>Midd2</i> | -12.33% | 1.12 | -2.53% | 0.43 | -12.01% | 1.56 | -12.65% | 1.51 |
| <i>Moebius</i> | -10.19% | 1.15 | -12.04% | 0.76 | -1.30% | 0.35 | -16.32% | 1.35 |
| <i>Teddy</i> | -8.02% | 0.92 | -2.93% | -0.51 | -7.35% | 0.87 | -8.14% | 0.92 |
| Average | -11.13% | 0.95 | -4.85% | 0.22 | -5.92% | 0.49 | -13.39% | 1.21 |

Table 1. Bitrate savings and averages $PSNR$ (dB) gains calculated against symmetric coding method using the Bjøntegaard measurement. Note that negative values correspond to bitrate savings compared to symmetric coding method, whereas for $PSNR$ the negative values indicate a reduction in $PSNR$

As shown in Figure 4, the SC method is the one required largest bitrate, thus providing a poor performance due to the fact that stereo pair is coded without any type of asymmetry. In contrast, the proposed method achieves similar quality while decreasing the required bitrate. The performance of the other asymmetric coding methods depend on the content of the images, they vary from one stereo pair to another. Overall, the AC_Down method achieves better performance than the two remaining asymmetric coding methods (AC_Blur and AC_Quality), but compared to the proposed method, it is clear that the latter is more efficient and outperforms the AC_Down method. For the AC_Blur method, despite the achieved bitrate saving, it is the one providing worst performance, since the AC_Blur method obtained the lowest *PSNR* values.

The proposed method provides significant bitrate saving without visual quality reduction of the reconstructed stereo images. This was achieved through the adaptation to the content of stereo pair and also thanks to the consideration of the quality of the other view. Unlike to the AC_Quality method, where the quality gap between view is fixed beforehand, thus not allowing any adaptation.

Bitrate savings and *PSNR* differences are given in Table 1. From the results of this Table, the bitrate saving achieved by the proposed method ranges from about 8.14% to 16.32% with an average of 13.39% compared with SC method. Also, AC_Down provides bitrate saving ranging from 5.64% to 15.31%, thus showing low performance than the proposed method. Finally, the weakest performance have been obtained by AC_Blur and AC_Quality methods, which achieved bitrate savings of 4.85% and 5.92% on average, respectively. In addition, the *PSNR* values obtained by the proposed method are much higher than the other methods. Moreover, for AC_Blur and AC_Quality methods in some instances, such as in Cloth3 and Teddy stereo pairs, we can note a *PSNR* degradation. Finally, these results clearly confirm the previous conclusions that the proposed method provides the best results.

According to the obtained results illustrated in Figure 4 and Table 1, thanks to the automatic and efficient adjusting of the quality gap between views, the proposed method provides a significant bitrate saving compared to the conventional coding approach without introducing noticeable distortion in 3D viewing.

4. Conclusion

In this paper, in order to enhance the asymmetric stereoscopic images coding, we proposed a novel model adjusting in automatic and adaptive manner the quality gap between images of the stereo pair. We derived a quality threshold through modeling the relationship between the quality of both views and their interview distortion and, also, thanks to the inclusion of HVS-inspired model. Based on this automatic thresholding of quality, the stereo pair is asymmetrically coded, thus reducing the required bitrate without causing visible artifacts in 3D viewing.

Experimental results showed that the proposed coding method can achieve considerable bitrate saving up to 16.32% compared to the conventional asymmetric coding methods based on a fixed threshold of asymmetry, which clearly demonstrates the effectiveness of the proposed coding approach.

In future work, the proposed method can be extended to the cases of stereoscopic/multiview videos, which constitutes new challenges, such maintaining the constant quality between successive frames.

References

- [1] Vetro, A., Tourapis, A. M., Müller, K., Chen T. (2011). 3D-TV Content Storage and Transmission, *IEEE Trans. Broadcast*, 57 (2) 384–394.
- [2] Wallace, G. K. (1992). The JPEG still picture compression standard, *IEEE Transactions on Consumer Electronics*, 38(1), xviii–xxxiv.
- [3] Acharya, T., Tsai, P. -S. (2005). JPEG2000 Standard for Image Compression, John Wiley & Sons, New Jersey.
- [4] Lukacs, M. E., (1986), Predictive coding of multi-viewpoint image sets, *In: Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP 86)*, p. 521–524.
- [5] Ouddane, S., Fezza, SA., Faraoun, K-M., (2013). Stereo Image Coding: State of the Art, in *Proc. IEEE International Workshop on Systems Signal Processing and their Applications (WoSSPA 2013)*, p. 122–126.

- [6] Moellenhoff, M. S., Maier, M. W., (1998). Transform coding of stereo image residuals, *IEEE Trans. on Image Processing*, 7 (6) 804–812.
- [7] Hachicha, W., Beghdadi, A., Cheikh, F. A., (2013). 1D directional DCT-based stereo residual compression, *In: Proc. European signal Processing Conference (EUSIPCO 2013)*.
- [8] Boulgouris, N. V., Strintzis, M. G., (2002). A family of wavelet-based stereo image coders, *IEEE Transactions on Circuits and Systems for Video Technology*, 12(10) 898–903.
- [9] Ellinas, J. N., Sangriotis, M. S., (2004). Stereo image compression using wavelet coefficients morphology, *Image and Vision Computing*, 22 (4) 281–290.
- [10] Frajka, T., Zeger, K., (2003). Residual image coding for stereo image compression, *Optical Engineering*, 42 (1) 182–189.
- [11] Maalouf, A., Larabi, M.-C., (2010). Bandelet-Based Stereo Image Coding, *In: Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010)*, Dalas Texas, USA, p. 698–701.
- [12] Woo, W., Ortega, A., (2000). Overlapped block disparity compensation with adaptive windows for stereo image coding, *IEEE Transactions on Circuits and Systems for Video Technology*, 10 (2) 194–200.
- [13] Bensalma, R., Larabi, M.-C., (2010). Optimizing the disparity map by the integration of HVS binocular properties for efficient coding of stereoscopic images, *In: Proc. IEEE 3DTV-Conference (3DTV-CON 210)*.
- [14] Ellinas, J. N., Sangriotis, M. S., (2006). Stereo Image Coder Based on the MRF Model for Disparity Compensation, *EURASIP Journal on Applied Signal Processing*, Article ID 73950, p. 1–13.
- [15] Stelmach, L. B., Tam, W. J., Meegan, D., Vincent, A., (2000). Stereo image quality: effects of mixed spatiotemporal resolution, *IEEE Trans. Circuits Syst. Video Technol.*, 10 (2) 188–193.
- [16] Aflaki, P., Hannuksela, M. M., Gabbouj, M., (2013). Subjective quality assessment of asymmetric stereoscopic 3-D video, *Journal of Signal, Image and Video Processing*, 9 (2) 331–345.
- [17] Seuntjens, P., Meesters, L., IJsselsteijn, W., (2006). Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation, *ACM Trans. Appl. Perception*, 3 (2) 95–109.
- [18] Saygili, G., Gurler, C. G., Tekalp, A. M., (2011). Evaluation of Asymmetric Stereo Video Coding and Rate Scaling for Adaptive 3D Video Streaming, *IEEE Trans. Broadcast.*, 57(2) 593–601.
- [19] Chang, Y., Kim, M., (2015). Binocular suppression-based stereoscopic video coding by joint rate control with KKT conditions for a hybrid video codec system, *IEEE Trans. Circuits Syst. Video Technol.*, 25(1) 99–111.
- [20] Fezza, SA., Larabi, M.-C., (2017). Perceptually Driven Nonuniform Asymmetric Coding of Stereoscopic 3D Video, *IEEE Trans. Circuits Syst. Video Technol.*, 27(10) 2231–2245.
- [21] Perkins, M. G., (1992). Data compression of stereopairs, *IEEE Trans. Commun.*, 40 (4) 684–696.
- [22] Aflaki, P., Hannuksela, M. M., Häkkinen, J., Lindroos, P., Gabbouj, M., (2010). Impact of downsampling ratio in mixed-resolution stereoscopic video, *In: Proc. 3DTV-Conference*.
- [23] De Silva, V. et al., (2012). Psycho-physical limits of interocular blur suppression and its application to asymmetric stereoscopic video delivery, *In Proc. 19th International Packet Video Workshop (PV 2012)*, p. 184–189.
- [24] Fehn, C., Kauff, P., Cho, S., Kwon, H., Hur, N., Kim, J. (2007). Asymmetric coding of stereoscopic video for transmission over T-DMB, *in Proc. 3DTV Conference*.
- [25] Chen, Y., Liu, S., Wang, Y.-K., Hannuksela, M. M., Li, H., Gabbouj, M., (2008). Low-complexity asymmetric multiview video coding, *In: Proc. International Conference on Multimedia and Expo (ICME 2008)*, p. 773–776.
- [26] Brust, H., Smolic, A., Müller, K., Tech, G., Wiegand, T., (2009). Mixed Resolution Coding of Stereoscopic Video for Mobile Devices, *In: Proc. 3DTV-Conference*.
- [27] Shao, F., Jiang, G., Wang, X., Yu, M., Chen, K., (2010). Stereoscopic video coding with asymmetric luminance and chrominance qualities, *IEEE Trans. Consum. Electron.*, 54 (6) 2460–2468.
- [28] Jayant, N., Johnston, J., Safranek, R., (1993). Signal compression based on models of human perception, *In: Proc. of the IEEE*, 81 (10) 1385–1422.

- [29] Zhao, Y., Chen, Z., Zhu, C., Tan, Y.-P., Yu, L., (2011). Binocular just-noticeable-difference model for stereoscopic images, *IEEE Signal Process. Lett.*, 18 (1) 19–22.
- [30] Hirschmüller, H., Scharstein, D., (2007). Evaluation of cost functions for stereo matching, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007).
- [31] Bjøntegaard, G., (2001). Calculation of average PSNR differences between RD-curves, in ITU-T SG 16 Q. 6 Video Coding Experts Group (VCEG), Document VCEG-M33, Austin, TX.