# ASYMMETRIC CODING USING BINOCULAR JUST NOTICEABLE DIFFERENCE AND DEPTH INFORMATION FOR STEREOSCOPIC 3D

Sid Ahmed FEZZA<sup>1</sup>, Mohamed-Chaker LARABI<sup>2</sup>, and Kamel Mohamed Faraoun<sup>1</sup>

<sup>1</sup>Department of Computer Science, Djillali Liabes University of Sidi Bel Abbes, Algeria <sup>2</sup>XLIM Laboratory, SIC Department, University of Poitiers, France

## ABSTRACT

The problem of determining the best level of asymmetry has been addressed by several recent works with the aim to guarantee an optimal binocular perception while keeping the minimum required information. To do so, subjective experiments have been conducted for the definition of an appropriate threshold. However, such an approach is lacking in terms of generalization because of the content variability. Moreover, using a fixed threshold does not allow an adaptation to the content and to the images' quality. The traditional asymmetric stereoscopic coding methods apply a uniform asymmetry by considering that all regions of an image have the same perceptual relevance which is not in compliance with the characteristics of human visual system (HVS). Consequently, this paper describes a fully automated model that dynamically determines the best bounds of asymmetry for each region of the image. Based on the Binocular Just Noticeable Difference (BJND) and the depth level in the scene, the proposed method achieves non-uniform reduction of spatial resolution of one view of the stereo pair with the aim to reduce bandwidth requirement. Experimental results show that the proposed method results in up to 43% of bitrate saving while outperforming the widely used asymmetric coding approaches in terms of 3D visual quality.

*Index Terms*— Asymmetric stereoscopic coding, binocular suppression, BJND, depth, HVS.

## 1. INTRODUCTION

The most widespread data format for 3D imaging is the stereoscopic technology [1,2]. Stereoscopic 3D (S3D) consists of capturing two images from two slightly different positions and provide each of them to each viewer's eye. Obviously, this results in doubling the required data for 2D for which appropriate compression schemes are very important.

A promising technique for coding stereo pairs known as asymmetric coding has focused considerable research efforts [4–17]. This approach is based on the so-called "binocular suppression theory" [3,4], which specifies that if one of the views of a given stereo-pair is altered, the 3D perceived quality will be close to the view with the highest quality [9], provided that the gap between both views does not exceed a threshold [5, 15, 16]. Depending on how the quality reduction of one of the views is achieved, the asymmetric stereoscopic coding methods can be classified into three categories: (*i*) spatial resolution reduction (spatial filtering) [7–14], (*ii*) asymmetric quantization (unequal QP) [5, 6, 15] and (*iii*) temporal resolution reduction [17]. With respect to the literature, the spatial filtering is by far the most explored and usually referred to as mixed resolution (MR) coding. In [16], authors compared spatial filtering and asymmetric quantization in a subjective experiment leading to the conclusion that the former outperforms the latter in terms of S3D quality. The third category has been explored in [4, 17] and it was observed that the asymmetry introduced by temporal resolution reduction produces perceptually noticeable artifacts, especially in case of high motion content.

MR coding was initially introduced by Perkins in [7] with the aim to reduce the bitrate for S3D delivery. Therefore, a spatial down-sampling of one the views is performed before encoding stage while the second view is kept as is. At the decoding stage, the view with lowest resolution is up-sampled to fit the resolution of the second view. Subjective tests have demonstrated that the perceived 3D fusion almost resembles to that obtained with original resolution image and several works have followed this concept since then [8–13].

Recent works have focused on identifying the limits of asymmetric coding or quantifying the just noticeable threshold of asymmetry between views of the stereo pair, for which the S3D effect is not affected [5, 14–16]. For instance, authors of [5] found that the just noticeable level of asymmetry depends on the display technology, *i.e.* about 9 dB for a parallax barrier display and 7 dB for a full resolution projection display. In the same vein, Shao *et al.* [15] showed that 2 dB of difference, between the left and right views, is the maximum tolerance level for which the viewer does not notice visual artifacts. Concerning the limit of asymmetric coding by reduction of spatial resolution, Aflaki *et al.* [14] made several trials with different downsampling ratios (1/2, 3/8, and 1/4) along both horizontal and vertical axes, and concluded that the range of downsampling ratios between 1/2 and 3/8 provides satisfactory 3D viewing experience.

All the previous works are based on subjective experiments and the determination of the just noticeable level of asymmetry highly relies on them. Unfortunately, it is difficult to generalize the derived threshold to any other stereo pair, because this threshold depends on the sequence content, display type, screen size and so on. A demonstration of this variability can be seen between [5] and [15], where, while both studies focused on asymmetric quantization following the same procedure, results gave very different threshold values. Another issue lies in the fact of assigning a single value to the threshold which does not allow any adaptation to the content and to the image original quality. Add to that, none of the attempts to determine the threshold have taken into account the view characteristics at full quality, as luminance or contrast, because for example, high contrast in one view may mask the visibility of impairments in the other.

In this paper, we propose a fully automated model that dynamically determines the best limits of asymmetry offering optimal 3D visual experience. The proposed method consists of adaptively selecting for each region of the image, the level of quality reduction. Our model is based on recent findings in visual perception; specifically the Binocular Just Noticeable Difference (BJND) [18] and the relationship between blur/sharpness and depth level [19]. We use the BJND model to determine the minimum distortions in one view that generate 3D perceptual difference, and the depth information to adjust the resolution.

The rest of the paper is organized as follows. Section 2 gives a brief overview of the BJND model. Section 3 provides details on the proposed approach. Experimental results are presented in section 4, including subjective tests and coding results. Finally, this paper ends with conclusions and some openings in relation to future work.

## 2. OVERVIEW OF THE BJND MODEL

Several studies in image/video processing have attempted to model the human visual system (HVS) features. Most of them used the well-known masking effect, usually known as Just Noticeable Difference (JND). Generally, the JND can be defined as the minimum change that could be noticed by a standard viewer. In order to build a Binocular-JND (BJND) model, Zhao *et al.* in [18] conducted psychophysical experiments. 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:

$$\begin{aligned} \text{BJND}_{l}(i,j) &= \text{BJND}_{l}(bg_{r}(i-d_{l},j), eh_{r}(i-d_{l},j), n_{r}(i-d_{l},j)) \\ &= A_{C}(bg_{r}(i-d_{l},j), eh_{r}(i-d_{l},j)) \\ &\times \left(1 - \left(\frac{n_{r}(i-d_{l},j)}{A_{C}(bg_{r}(i-d_{l},j), eh_{r}(i-d_{l},j))}\right)^{\lambda}\right)^{\frac{1}{\lambda}} \end{aligned}$$
(1)

where *i* and *j* are the pixel coordinates,  $d_l$  is the horizontal disparity value at pixel (i, j). The parameter  $\lambda$  controls the influence of the noise in the right image, and it was suggested in [18] that  $\lambda = 1.25$ . It should be noted that BJND<sub>l</sub> 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 (found by stereo matching). Where no noise in the right image, *i.e.*  $n_r(i-d, j) = 0$ , the BJND<sub>l</sub> is reduced to  $A_C$ , which is defined as

$$A_C(bg, eh) = A_{limit}(bg) + K(bg) \cdot eh \tag{2}$$

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

$$A_{limit}(bg) = \begin{cases} 0.0027 \cdot (bg^2 - 96 \cdot bg) + 8, & \text{if } 0 \le bg < 48\\ 0.0001 \cdot (bg^2 - 32 \cdot bg) + 1.7, & \text{if } 48 \le bg \le 255 \end{cases}$$
(3)

$$K(bg) = -10^{-6} \cdot (0.7 \cdot bg^2 + 32 \cdot bg) + 0.07 \tag{4}$$

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

$$eh(i,j) = \sqrt{E_H^2(i,j) + E_V^2(i,j)},$$
 (5)

$$E_k(i,j) = \frac{1}{24} \sum_{h=1}^{5} \sum_{v=1}^{5} I(i-3+h, j-3+v) \cdot G_k(h,v), \quad (6)$$
  
with  $k = H, V.$ 

where I(i, j) represents the luminance value of pixel (i, j), and the detailed representation of the Sobel operator  $G_k(h, v)$  can be found in [18].

## 3. PROPOSED ADAPTIVE ASYMMETRIC STEREOSCOPIC CODING

## 3.1. Motivations

As stated before, when the left and right images are presented to the viewer with different levels of sharpness, *i.e.* one image of the stereo pair is more blurry, the resulting 3D perception is close to the sharper image. On the other hand, the HVS tolerates a certain limited level of blur in one view that does not impair the 3D viewing experience. The principle of the proposed asymmetric stereoscopic coding approach relies on the fact that, the blur introduced in one view of a stereo pair should lead to a change below the binocular visibility threshold (BJND), thus avoiding visible artifacts. Moreover, as shown by subjective experiments described in [19], the sharpness visibility at different depth levels is different. The perceived blur of the object is highly dependent on the objects' distance (position in the scene), *i.e.* the closer object appears sharper than deeper objects. Accordingly, we use the depth information to control the strength of the introduced blur. In traditional methods of MR coding, one view is fully blurred at a fixed level of blur. However, the binocular visual properties are not spread equally across the stereoscopic images, because the response of the HVS to distortion in different regions is not uniform [20]. For example, the occluded regions cannot be binocularly compensated (masked), because they have no corresponding region in the other view. Thus, the blur effect in such regions may be noticed. Therefore, we propose a novel MR coding that enhances the concept of asymmetric coding relying on several aspects of the HVS such as the BJND model and the relationship between blur/sharpness and depth level. The proposed model adjusts the resolution dynamically to adapt to each feature of a given region in the image.

#### 3.2. Filter Design

Usually, there are two ways to achieve asymmetric blurring in MR coding. The first one is downsampling an image to a smaller size, and the second is applying blur filter (*e.g.* disk filter, Gaussian filter, ...). From the visual quality standpoint, both procedures have the same visual effect of blurring. However, downsampling and then upsampling back for 3D visualization, introduces more visual artifacts than using blur filter [11,12]. Also, from a compression point of view, combining inter-view prediction with downsampling concept requires an extra buffer size, in addition to design changes affecting low levels of encoding and decoding processes [8, 10, 13]. Hence, the complexity of coding and decoding is increased. In contrast, the disk blur filter is applied as preprocess step keeping the coding and decoding unchanged. Finally, using a blur filter allows to assign different degrees of blur to different regions of the image, thus providing more flexibility.

In our model, we use a circular kernel filter as blur kernel instead of the well-known Gaussian filter. In order to tune the strength of the blur filter, the disk filter depends on a single parameter, which is the diameter of the filter (kernel width), in contrast to Gaussian filter that depends on the standard deviation ( $\sigma$ ) as well as the kernel width. Consequently, with a disk filter, we have fewer variables to optimize. Add to that, disk filter offers a radial symmetry as well as a flat response [11].

#### 3.3. Proposed Asymmetric Coding Method

Given a stereo pair of images, denoted by  $I_l$  for the left view and  $I_r$  for the right. In the following, we consider the right image as the sharper image and it is not altered, while the left image is the image to be blurred (this choice is arbitrary but may rely on the notion of eye dominance). The aim is to reduce the spatial resolution of the left image, in order to decrease the bitrate without causing visible artifacts in 3D viewing. To achieve this, we blur  $I_l$  by a disk filter F defined as follows:

$$\tilde{I}_{l}^{(R_{i,j})}(i,j) = I_{l}(i,j) * F_{R_{i,j}}$$
(7)  
with  $R_{i,j} = \underset{k>0}{\arg\max} g_{k}(i,j)$   
 $= \{k \mid |I_{l}(i,j) - \tilde{I}_{l}^{(k)}(i,j)| < \delta_{i,j} \cdot VT_{i,j}\}$ 

where \* is a convolution product, and  $\delta_{i,j}$  is the weighting factor related to the stereo matching result (disparity map), the definition of this factor will be discussed later.  $I_l(i, j)$  denotes the luminance value of the pixel at position (i, j),  $R_{i,j}$  is diameter of the disk filter  $F_{R_{i,j}}$  that determines the strength of blur for each pixel (i, j), and  $\tilde{I}_l^{(k)}$  is the output of the filtered left image with a k diameter disk filter. The change introduced by the blur shall not exceed the visual threshold  $VT_{i,j}$  defined by equation (8):

$$VT_{i,j} = \frac{D_{max}}{d_l(i,j)} \cdot \text{BJND}_l(i,j)$$
(8)

where  $d_l(i, j)$  is the disparity value of pixel (i, j), and  $D_{max}$  is the maximum value of the disparity map.  $VT_{i,j}$  depends on two factors: the BJND<sub>l</sub>(i, j) which represents the Binocular-JND (described in Section 2) of the left image at pixel position (i, j), and the depth information.

It is known that the disparity is inversely proportional to depth, meaning that objects with greater disparity are closest, and vice versa. Accordingly,  $VT_{i,j}$  of closer pixels  $(d_l(i, j) \cong D_{max})$  are mainly determined by BJND<sub>l</sub>, unlike to deeper pixels which are more blurred, because  $VT_{i,j}$  is higher. This process is applied only to the luminance component, to all pixels of the left view, except those with zero disparity (*e.g.* occluded pixels or bad matches). For the latter, they are kept unaltered. In our implementation, to estimate the disparity map, we use the stereo matching algorithm described in [21].

Since  $VT_{i,j}$  is dependent on two factors which in turn are based on the disparity map, and as in practice the stereo matching algorithms can cause some bad matches, we introduce a weighting factor  $\delta_{i,j}$  to control the influence of  $VT_{i,j}$  according to the reliability of the stereo matching result. To achieve this, we use a dissimilarity function that assess the stereo matching result in each pixel as follows:

$$\delta_{i,j} = a_1 exp(-a_2 \cdot |ZNCC(i,j,d_l)|) \tag{9}$$

where  $a_1$  and  $a_2$  are two constants determined experimentally, and ZNCC (Zero mean Normalized Cross Correlation) is windowbased dissimilarity measures defined by equation (10). In the latter equation, W denotes a  $5 \times 5$  window centered at pixel (i, j), and  $\overline{I}_*(i, j)$  is the mean value computed over all pixels inside W. The ZNCC ranges from [-1, 1], and it is used in equation (9) to exponentially decrease the influence of  $VT_{i,j}$  according to the quality of the disparity map. On the other hand, if the disparity value is not reliable, then ZNCC is high, and consequently, the value of the pixel is not significantly changed by the proposed process.

## 4. EXPERIMENTAL RESULTS

In this section, simulation results are provided to evaluate the performance of the proposed asymmetric stereoscopic coding. Simulations have been carried out on the Middleburry stereo dataset [22]. The used test images were Art, Book, Cloth3, Cones, Midd1, Midd2, Moebius, Reindeer and Teddy. The objective and subjective performance of the proposed adaptive asymmetric stereoscopic coding method and symmetric stereoscopic coding method (denoted as SSC) as well as traditional uniform MR coding (UMR) methods were compared. Specifically, we considered two traditional uniform asymmetric stereoscopic coding methods. The first one achieves the MR coding by applying downsampling ratio of 1/2 along both coordinate axes (denoted as UMR-D), and the second consists in applying disk filter with a fixed diameter of 2 (denoted as UMR-B). The values of 1/2 and 2 have been demonstrated experimentally in [14] and [11], respectively, so that they provide satisfactory 3D quality. Due to page limitation, results will be given only for some of the images and mean values are provided when possible.

#### 4.1. Objective performance evaluation

First, we demonstrate the effectiveness of the proposed method in terms of coding performance. To this end, we used the JMVM (Joint Multi-view Video Model) software [24], which is the multi-view video coding (MVC) reference software based on H.264/AVC [23], with different QP values (QP=22, 27, 32, 37). Usually, H.264/MVC is used for coding multi-view video but in our case it is used to enable the inter-view prediction. The right image was coded as I frame (reference image) and the left image (target image) was predicted from the reference one, usually this type of image is referred as P frame. To allow the inter-view prediction in UMR-D method, the right image is downsampled by 1/2 (along both coordinate axes) at the reference picture buffer level. Finally, after decoding, the downsampled left image is upsampled for PSNR calculation. Table 1 lists the bitrate saving and average PSNR results generated using the Bjøntegaard measurement [25] for some images. It also provides the average values for the whole test-set.

The results show that the proposed method can achieve bitrate saving ranging from 34% to 43% compared with SSC method, also UMR-B and UMR-D with lower performance than the proposed model, provide up to 41.47% and 41.67% bitrate saving, respectively. These results can be explained by the fact that some regions are too blurred achieving thus more reducing bitrate compared to uniform blur. Moreover, regions with less blur greatly benefit from

$$ZNCC(i, j, d_l) = \frac{\sum_{(i, j) \in W} (I_l(i, j) - \bar{I}_l(i, j)) \cdot (I_r(i - d_l, j) - \bar{I}_r(i - d_l, j))}{\sqrt{\sum_{(i, j) \in W} (I_l(i, j) - \bar{I}_l(i, j))^2 \cdot \sum_{(i, j) \in W} (I_r(i - d_l, j) - \bar{I}_r(i - d_l, j))^2}}$$
(10)

Method	UMR-B		UMR-D		Proposed	
Image	Bitrate saving	PSNR	Bitrate saving	PSNR	Bitrate saving	PSNR
Art	35.69%	2.95	33.90%	2.75	36.75%	3.04
Books	38.70%	2.79	36.37%	2.50	38.04%	2.93
Moebius	34.61%	2.01	23.38%	1.25	34.62%	2.03
Cloth3	40.87%	3.94	40.78%	3.52	40.92%	4.22
Average	39,97%	3,32	37,63%	3,05	41,17%	3,58

Table 1: Bitrate saving and average PSNR (dB) comparisons measured against SSC method using Bjøntegaard measurement [25].

the inter-view prediction than the others UMR methods. In terms of average PSNR, Table 1 shows that the proposed method outperforms the others for all displayed images.



Fig. 1: Rate/distortion curves for left image of Cones stereo pair.

In addition to the Bjøntegaard measurement, Figure 1 provides rate/distortion curves only for the left view of the stereo pair *Cones* (the right being the reference as stated previously). It clearly confirms the previous conclusions since our proposal provides the best results.

#### 4.2. Subjective performance evaluation

In order to confirm the good results obtained with our approach, we propose a subjective validation so as to quantify the 3D visual experience of the viewers. Therefore, the test was performed in the XLIM lab test-room, calibrated following the ITU-R BT.500 recommendations [26], using a Hyundai S456D, a passive film pattern retarder stereoscopic 3D TV. The display was placed 0.5 m from the back wall and 4H from the viewer (H being the heigh of the image). Each viewer adjusted the height of their chair so that the position of his/her eyes were at about the same as the height of the center of the display. The room lighting was controlled and the display has been calibrated using the i1display2 from Gretag Macbeth<sup>(R)</sup>.

Twelve naive observers with age raging from 25 to 35 participated to the test. They have been pre-screened for visual acuity, color blindness using Ishihara test and depth blindness using Randot Stereo test.

A set of 36 stereoscopic images, corresponding to 9 original pairs with associated results obtained using UMR-B, UMR-D and our proposal, was randomized and presented sequentially. We opted



**Fig. 2**: DMOS values for each image of the test-set using UMR-B, UMR-D and proposed approach.

for an ACR-5 (Absolute Category Rating) method with hidden reference [26] for this test. The aim being in close to real condition when someone use his display to view a stereo-pair without any cue about the original. Observers were asked to rate each S3D image on a quality scale composed of 5 categories (Bad, Poor, Fair, Good, Excellent). Before the experiment starts, observers were asked to read the instructions explaining the task before going to the training session.

In order to account for the influence of the hidden reference, the final scores DMOS (Difference Mean Opinion Score) are calculated using the following equation:

$$DMOS_{test} = MOS_{test} - MOS_{Original} + 5$$
(11)

where  $MOS_{test}$  and  $MOS_{Original}$  are respectively the score of a given stereo pair for one of the used method and the score of the original image.

Figure 2 gives the DMOS for all images from the test-set. Values above 5 indicate that observers scored the image better than the original one. One can notice that the proposed approach obtains very good results almost for all images except *Moebius*. This confirms the results obtained from the objective study.

## 5. CONCLUSION AND FUTURE WORKS

We proposed in this paper a novel asymmetric stereoscopic coding method. The proposed approach enhances the MR coding concept by introducing a non-uniform reduction of spatial resolution allowing decreasing the bandwidth required for S3D delivery. The proposed method determines automatically and adaptively the maximum tolerance level of blur for each region (pixel) based on the BJND and the depth level, making the blur effect almost transparent to the viewer. Objective and Subjective performance evaluation results have demonstrated that the proposed approach can provide significant bitrate saving without noticeable visual quality losses. From a visual quality standpoint, the produced image quality shows a gap compared to the state-of-the-art methods.

As future directions, we plan to extend the proposed method to stereo video coding applications by taking into account the temporal properties of the HVS. Another work that deserves to be addressed is to preprocessing of the disparity map, in order to assign a valid value to the bad matches and occluded regions.

### 6. REFERENCES

- A. Vetro, A. M. Tourapis, K. Müller, and T. Chen, "3D-TV Content Storage and Transmission," *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 384-394, Jan. 2011.
- [2] A. Smolic, K. Müller, P. Merkle, P. Kauff, and T. Wiegand, "An Overview of Available and Emerging 3D Video Formats and Depth Enhanced Stereo as Efficient Generic Solution," in *Proc. Picture Coding Symposium (PCS 2009)*, Chicago, IL, USA, May. 2009, pp. 1-4.
- [3] B. Julesz, Foundations of Cyclopean Perception. Univ. Chicago Press, 1971.
- [4] L. B. Stelmach, W. J. Tam, D. Meegan, and A. Vincent, "Stereo image quality: effects of mixed spatio-temporal resolution," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 2, pp. 188-193, Mar. 2000.
- [5] G. Saygili, C. G. Gurler, and A. M. Tekalp, "Evaluation of Asymmetric Stereo Video Coding and Rate Scaling for Adaptive 3D Video Streaming," *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 593-601, Jun. 2011.
- [6] L. B. Stelmach and W. J. Tam, "Stereoscopic image coding: Effect of disparate image-quality in left-and right-eye views," *Signal Process. Image Commun.*, vol. 14, no. 1, pp. 111-117, Nov. 1998.
- [7] M. G. Perkins, "Data compression of stereopairs," *IEEE Trans*actions on Communications, vol. 40, no. 4, pp. 684-696, Apr. 1992.
- [8] C. Fehn, P. Kauff, S. Cho, H. Kwon, N. Hur, and J. Kim, "Asymmetric coding of stereoscopic video for transmission over T-DMB," in *Proc. 3DTV Conference*, Kos Island, Greece, May. 2007.
- [9] P. Aflaki, M. M. Hannuksela, J. Häkkinen, P. Lindroos, and M. Gabbouj, "Subjective study on compressed asymmetric stereoscopic video," in *Int. Conf. Image Process. (ICIP 2010)*, Hongkong, China, Sep. 2010, pp. 4021-4024.
- [10] Y. Chen, S. Liu, Y.-K. Wang, M. M. Hannuksela, H. Li, and M. Gabbouj, "Low-complexity asymmetric multiview video coding," in *Proc. International Conference on Multimedia and Expo (ICME 2008)*, Hannover, Germany, Apr. 2008, pp. 773-776.
- [11] A. K. Jain, A. E. Robinson, and T. Q. Nguyen, "Comparing Perceived Quality and Fatigue for Two Methods of Mixed Resolution Stereoscopic Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol.xx, no.xx, Aug. 2013.
- [12] A. E. Robinson, A. K. Jain, M. Scott, D. MacLeod, and T. Q. Nguyen, "Apparent sharpness of 3D video when one eye's view is more blurry," *i-Perception*, vol. 4, no. 6, pp. 456-467, 2013.
- [13] H. Brust, G. Tech, K. Müller, and T. Wiegand, "Mixed resolution coding with inter view prediction for mobile 3DTV," in *Proc. 3DTV-Conference*, Tampere, Finland, Jun. 2010.
- [14] P. Aflaki, M. M. Hannuksela, J. Häkkinen, P. Lindroos, and M. Gabbouj, "Impact of downsampling ratio in mixed-resolution stereoscopic video," in *Proc. 3DTV-Conference*, Tampere, Finland, Jun. 2010.
- [15] F. Shao, G. Jiang, X. Wang, M. Yu, and K. Chen, "Stereoscopic video coding with asymmetric luminance and chrominance qualities," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 6, pp. 2460-2468, Nov. 2010.

- [16] V. De Silva, H. Arachchi, E. Ekmekcioglu, A. Fernando, S. Dogan, A. Kondoz, and S. Savas, "Psycho-physical limits of interocular blur suppression and its application to asymmetric stereoscopic video delivery," in *Proc. 19th International Packet Video Workshop (PV 2012)*, Munich, Germany, May. 2012, pp. 184-189.
- [17] A. Aksay, C. Bilen, E. Kurutepe, T. Ozcelebi, G. B. Akar, R. Civanlar, and A. M. Tekalp, "Temporal and spatial scaling for stereoscopic video compression," in *Proc. European Signal Processing Conference (EUSIPCO 2006)*, Florence, Italy, Sep. 2006.
- [18] Y. Zhao, Z. Chen, C. Zhu, Y.-P. Tan, and L. Yu, "Binocular just-noticeable-difference model for stereoscopic images," *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 19-22, Jan. 2011.
- [19] M. M. Subedar and L. J. Karam, "Increased depth perception with sharpness enhancement for stereo video," in *Proc. SPIE Stereoscopic Displays and Applications XXI*, vol. 7524, San Jose, CA, USA, Feb. 2010.
- [20] F. Shao, W. Lin, S. Gu, G. Jiang, and T. Srikanthan, "Perceptual Full-Reference Quality Assessment of Stereoscopic Images by Considering Binocular Visual Characteristics," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1940-1953, May. 2013.
- [21] A. Klaus, M. Sormann, and K. Karner, "Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure," in *Proc. Int. Conf. Pattern Recognition (ICPR* 2006), Hong Kong, Aug. 2006, pp. 15-18.
- [22] Middleburry stereo vision website, http://vision.middlebury.edu/stereo.
- [23] A. Vetro, T. Wiegand, and G. J. Sullivan, "Overview of the Stereo and Multiview Video Coding Extensions of the H.264/AVC Standard," *Proceedings of the IEEE*, 2011.
- [24] A. Vetro, P. Pandit, H. Kimata, A. Smolic, and Y. K. Wang, "Joint Multiview Video Model (JMVM) 8.0," ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16, Doc. JVT-AA207, Apr. 2008.
- [25] G. Bjøntegaard, "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, Apr. 2-4, 2001.
- [26] ITU-R Rec. BT.500, Methodology for the subjective assessment of the quality of television pictures", 46 pages, Geneva, Switzerland", 2012.