STEREO IMAGE QUALITY ASSESSMENT USING A BINOCULAR JUST NOTICEABLE DIFFERENCE MODEL

Walid Hachicha [1], *Azeddine Beghdadi* [1], *Faouzi Alaya Cheikh* [2]

[1]L2TI, Institut Galilée, Université Paris 13, Sorbonne Paris Cité, France
[2]The Norwegian Colour and Visual Computing Lab, Gjøvik University College, Norway
walid.hachicha@univ-paris13.fr, azeddine.beghdadi@univ-paris13.fr, faouzi.cheikh@hig.no

ABSTRACT

This paper presents a novel full-reference Stereo Image Quality Assessment (SIQA) measure based on well understood characteristics of the human visual system (HVS), namely contrast sensitivity and frequency and directional selectivity. Additionally, the proposed metric takes into account the stereo interplay between the two views, where one view may affect our perception of the overall quality of the stereo image pair. Therefore, a Binocular Just Noticeable Difference (BJND) model is used to compute the distortion visibility threshold, and the binocular suppression theory is considered in the proposed metric. The scored 3D LIVE IQA database is used to evaluate the correlation of the proposed metric with the DMOS subjective score provided by the database. The obtained experimental results show that the proposed metric correlates much better with the DMOS score than the state-of-the-art metrics do.

***Index Terms*—** Stereo image quality assessment, HVS model, BJND, Binocular suppression.

1. INTRODUCTION

We are witnessing a rapid development of stereo and multi-view systems and a wide adoption of these systems. Therefore, Stereoscopic Image Quality Assessment (SIQA) has become an important and challenging problem faced in numerous applications such as 3D acquisition, stereo visualization, stereo compression and 3D enhancement. Objective image quality assessment plays an important role in many image processing applications. It may serve as a decision support tool for selecting the algorithm or the system that provides the best image quality. It may also be embedded in an image processing framework to generate optimal algorithm parameters. Rate/Distortion optimization approach for image compression is one of such potential applications where objective image quality plays a prominent role.

For applications where a human viewer is the ultimate receiver, the best way to assess the image quality is through a subjective measurement. The perceptual human judgment is frequently expressed in terms of Mean Opinion Score (MOS). Subjective evaluation however, is time-consuming, costly, depends upon viewer's physical condition, emotional state, personal experience, context of preceding display and thus is highly inconvenient in most applications. In addition, subjective measures cannot be incorporated into automatic image processing systems especially for real-time applications. Furthermore, when it comes to stereo images, sometime objective measures may become unpredictable due to additional human physical factors related to stereovision such as fatigue and visual discomfort. Therefore, it is crucial to develop reliable quantitative measures that can gauge automatically the perceived image quality. The robustness of such measures is evaluated by how well they correlate with human observers' judgment.

Many Image Quality Assessment (IQA) metrics have been proposed for 2D image to quantify objectively the image quality. Existing methods can be roughly classified into three categories, i.e. full-reference (FR), reduced-reference (RR) and no-reference (NR) metrics. FR approaches need a complete reference image, while RR ones need only parts (or features) of the reference image. The NR ones do not require any information about the original image and are typically designed for specific distortions, for more details reader may refer to the survey in [1]. It is worth noting that this classification is also valid for stereoscopic 3D images. In order to develop good objective perceptual stereo quality metrics, it is paramount to understand the human perception of 3D quality. Human perception of 3D image quality differs from that of the 2D quality perception because additional binocular cues have to be considered in the 3D case. The most important one is the depth perception and how it contributes to the overall perceived 3D image quality.

The objective of this paper is to design a FR metric based on human visual perception findings, especially contrast masking effect [2], and Binocular Just Noticeable Difference (BJND) [3]. Furthermore, in this work the binocular suppression is modelled and incorporated into the proposed metric.

The rest of this paper is organized as follows. In section 2, we start by a brief review of SIQA state-of-the-art. Section 3 gives an overview of the BJND model. Section 4 describes the proposed metric. In section 5, we present the experimental results. Finally, conclusions and possible further research directions are presented in section 6.

2. SIQA STATE-OF-THE-ART

In the literature many efforts have been devoted to designing SIQA metrics [4–14]. In [4], the authors studied the use of some well-known 2D objective metrics for quality assessment of stereo images. Their study consisted in combining the two scores for the right and the left images using three approaches, i.e., average, main eye and visual acuity. No information about the depth perception was taken into account however. Benoit et al. [5] proposed a metric combining two measures. The first one is the difference between original (left or right) and distorted image. The second is the difference between disparity maps before and after applying distortion to the stereo pair. For the combination operation, both global and local approaches were proposed. A similar approach was adopted by You et al. [7], who investigated the combination of eleven 2D image quality metrics with the disparity distortion. Thus these metrics depended on the disparity estimation method and its robustness against distortion. In [8] a metric based on the matching of regions with high spatial frequency in the left and right views is proposed for evaluating the quality of compressed images. In [9], Sazzad et al. proposed a NR metric for JPEG coded stereoscopic images based on local features, i.e. blockiness and zero-crossing within the edge area, flat and texture as well as disparity information of plane and non-plane areas. A stereo sense assessment metric (SSA) based on the disparity distribution has been proposed in [11], where only absolute disparity was used. In [12], authors proposed a perceptual model consisting of wavelet-based perceptual decomposition and contrast conversion and masking. Hewage et al. [13] proposed RR quality metric for depth map transmission based on edge detection. In [14], authors proposed metric based on the computation of the cyclopean images of the reference and distorted pairs. The properties of the HVS are integrated in the metric by performing a Contrast Sensitivity Function (CSF) filtering of the cyclop images after a wavelet decomposition and rational thresholding to obtain the sensitivity coefficients. The metric is the average of disparities coherence and perceptual difference between the reference and distorted cyclop images.

3. BJND MODEL

The BJND model measures the perceptible distortion threshold of binocular vision for stereoscopic images. In [3], it has been demonstrated, through many psychophysical experiments, that the BJND depends on two of the HVS characteristics: luminance adaptation and binocular contrast masking. Given the left and right images the BJND at the left view, i.e., BJND_l, is defined as follows [3]:

$$\begin{aligned}\text{BJND}_l(i,j,d_l) &= A_C(bg_r(i-d_l,j), eh_r(i-d_l,j)) \\ &\times (1-(\frac{n_r(i-d_l,j)}{A_C(bg_r(i-d_l,j), eh_r(i-d_l,j))})^{\lambda})^{\frac{1}{\lambda}}, \end{aligned} \quad (1)$$

where A_C is the elevated threshold function due to the contrast masking effect, d_l is the disparity value at pixel coordinate (i,j), note that the parallel camera model is assumed here, thus the disparity value represents the horizontal displacement of the pixel from left to right views. The exponent λ allows adjusting the noise influence in the right view, and according to the psychophysical experiment it is set to 1.25. We note that the BJND_l depends on the background luminance level bg_r, the edge height eh_r and the noise amplitude n_r of the corresponding pixel in the right view. For a noise-free view the BJND_l is reduced to the A_C term defined as follows:

$$A_C(bg, eh) = A_{limit}(bg) + B(bg).eh \quad (2)$$

The A_{limit} and B functions are derived using psychophysical experiment and defined as follows :

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

$$B(bg) = -10^{-6}(0.7bg^2 + 32bg) + 0.07. \quad (4)$$

The background luminance bg is computed by averaging the luminance values over a 5×5 window centred at the corresponding pixel position and the edge height eh is given by the 5×5 Sobel operators as follows:

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

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

where $k \in \{H, V\}$, $I(i,j)$ denotes the luminance value at (i,j) and

$$G_H = \begin{pmatrix} -1 & -2 & 0 & 2 & 1 \\ -2 & -3 & 0 & 3 & 2 \\ -3 & -5 & 0 & 5 & 3 \\ -2 & -3 & 0 & 3 & 2 \\ -1 & -2 & 0 & 2 & 1 \end{pmatrix}, G_V = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 2 & 3 & 5 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -3 & -5 & -3 & -2 \\ -1 & -2 & -3 & -2 & -1 \end{pmatrix}.$$

4. THE PROPOSED METRIC

The proposed metric is based on the fact that, the degradation is perceived only when it yields a change above the binocular visibility threshold. It is also sensitive to the influence of one view on the perception of the other (binocular suppression), i.e. a high contrasted area in one view may mask the visibility of degradation in the corresponding region in the other view. The proposed stereo image quality prediction model is composed of three stages. First, perceptual images of original and degraded stereo pairs are computed. Second, the BJNDs for both reference and distorted stereo pairs are estimated. Finally, a pooling step to model the global distortion metric is performed.

The first stage aims at computing a perceptual representation for each view of reference and degraded stereo pairs. The proposed algorithm builds on the work done by Daly [2]. For each view $E\times F$+1 images are built as follows:

$$\begin{aligned} m_{e,f}(i,j) &= \mathcal{F}^{-1}(\mathcal{F}(I)(u,v).\text{CSF}(u,v).\text{Cortex}_{e,f}(u,v)), \\ m_E(i,j) &= \mathcal{F}^{-1}(\mathcal{F}(I)(u,v).\text{CSF}(u,v).\text{base}(u,v)), \end{aligned} \quad (7)$$

where I is the input luminance image processed by the amplitude non-linearity because perception of lightness is a nonlinear function of luminance, CSF is the 2D contrast sensitivity function, $\text{Cortex}_{e,f}$ is the (e,f) cortex band, base is the lowest frequency filter and $\mathcal{F}$ stands for the Fourier transform.

The CSF describes the variation of visual sensitivity as a function of spatial frequencies and the $\text{Cortex}_{e,f}$ describes a band decomposition filter where e is the radial and f is the orientation frequencies indexes, $e \in \{0,..,E-1\}$ and $f \in \{0,..,F-1\}$ (E and F are respectively the number of radial bands and the number of orientations). These two terms of the perceptual representation account for the frequency and directional selectivity of the HVS. More details on the models used for amplitude non-linearity, the CSF and the Cortex could be found in [2].

Let $\mathcal{A}$ be a filter set including cortex band filters and the base band filter. Once the filtering process done, we compute for each filter α the local variation of the differential visibility threshold elevation defined by the following equation:

$$\tilde{m}_\alpha(i,j) = (1 + (k_1(k_2|m_\alpha(i,j)|)^s)^b)^{\frac{1}{b}}, \quad (8)$$

where $k_1 = 0.0153$, $k_2 = 392.5$, $b = 4$ and $s \in [0.6, 1]$. Finally, we define the resulting perceptual image by retaining the maximum visibility value of each pixel, i.e. the output image is defined as follows:

$$m(i,j) = \max(\tilde{m}_\alpha(i,j))_{\alpha\in\mathcal{A}}. \quad (9)$$

The next step is the BJND model computation, which is an important task in the proposed algorithm. The BJND models of the reference and distorted stereo pairs are computed after disparity estimation. Disparity estimation determines the pixel displacement between the two views of a stereo pair. In our implementation, disparity map can be estimated using one of the different methods, based on the Markov Random Field (MRF) model [15], provided by the software available in [16].

Before deriving the final stereo quality metric, pixels in the reference stereo pair and their corresponding pixels in the degraded one are segmented into four disjoint classes for each view, $\mathcal{O}_v$, $\mathcal{T}_v$, $\mathcal{S}_v$ and $\mathcal{R}_v$, $v \in \{\text{left}, \text{right}\}$. In the following, left and right views are referred to l and r and $\tilde{v}$ refers to a degraded view.

- *Occlusion class* ($\mathcal{O}_l$): contains occluded pixels in the left view, identified by a zero disparity value and/or correponds to the overflowed disparity-shifted pixels.
- *Invisible distortion class* ($\mathcal{T}_l$): contains non-occluded pixels having luminance distortion lower than its visibility threshold BJND_l in which the change is detected by the majority of observers (threshold visibility criteria).
- *Binocular suppression class* ($\mathcal{S}_l$): pixels belonging to this class are non-occluded and satisfy the binocular suppression criteria. First, the local contrast around a left view degraded pixel should be higher than its corresponding pixel in the degraded right view (left to right local contrast comparison criteria). Second, the inter-difference between the two views at this pixel should be less than the visibility threshold $\text{BJND}_{\tilde{l}}$ of the degraded left view. Pixels should not satisfy the threshold visibility criteria.
- *Binocular rivalry class* ($\mathcal{R}_l$): contains non-occluded pixels satisfying the left to right local contrast comparison criteria and the inter-difference between the two views at these pixels exceed the visibility threshold $\text{BJND}_{\tilde{l}}$. The threshold visibility criteria is not respected for this class.

Mathematically, $\mathcal{O}_v$, $\mathcal{T}_v$, $\mathcal{S}_v$ and $\mathcal{R}_v$ are defined as follows:

$$\mathcal{O}_l = \{(i,j) \in l, d_l(i,j) = 0\} \cup \{(i,j) \in l, i\text{-}d_l < 0\},$$

$$\mathcal{T}_l = \{(i,j) \in \bar{\mathcal{O}}_l, \underbrace{\sum_{(p,q)\in\mathcal{B}} |I_l(p,q)\text{-}I_{\tilde{l}}(p,q)| < \sum_{(p,q)\in\mathcal{B}} \text{BJND}_l(p,q)}_{\text{threshold visibility criteria}}\},$$

$$\mathcal{S}_l = \bar{\mathcal{O}}_l \cap \bar{\mathcal{T}}_l \cap \mathcal{C}_l \cap \mathcal{I}_l, \quad \mathcal{R}_l = \bar{\mathcal{O}}_l \cap \bar{\mathcal{T}}_l \cap \mathcal{C}_l \cap \bar{\mathcal{I}}_l$$

$$\mathcal{C}_l = \{(i,j) \in l, \underbrace{\sum_{(p,q)\in\mathcal{B}} C_{\tilde{l}}(p,q) > \sum_{(p,q)\in\mathcal{B}} C_{\tilde{r}}(p\text{-}d_l,q)}_{\text{left to right local contrast comparison criteria}}\},$$

$$\mathcal{I}_l = \{(i,j) \in \bar{\mathcal{O}}_l, \underbrace{\sum_{(p,q)\in\mathcal{B}} |I_{\tilde{l}}(p,q)\text{-}I_{\tilde{r}}(p\text{-}d_l,q)| < \sum_{(p,q)\in\mathcal{B}} \text{BJND}_{\tilde{l}}(p,q)}_{\text{Inter-difference threshold visibility criteria}}\},$$

where $C_v, v \in \{\tilde{l}, \tilde{r}\}$ is the local contrast computed as in [17], $\mathcal{B}$ is a square block of size $\omega \times \omega$ centered at pixel (i,j) (occluded pixels are not considered), d_l is the disparity value at (i,j) of the left view. d_r, $\mathcal{O}_r$, $\mathcal{T}_r$, $\mathcal{S}_r$, $\mathcal{R}_r$, $\mathcal{I}_r$ and $\mathcal{C}_r$ are defined in similar fashion by substituting l by r and the horizontal pixel position is positive shifted, i.e. $i + d_r$. Note that $i + d_r$ should be less than the image width size, otherwise, pixel at (i,j) position of the right view is considered as occluded right.

The inter-difference threshold visibility and the side-to-side local contrast comparison can model the binocular suppression theory due to the fact that local high contrasted area in one side tends to suppress those in the other side with low contrast. If the inter-view distortion is visible, i.e., it exceeds the BJND, binocular rivalry may occur.

Finally, the stereo quality metric is derived as following:

$$SM = \sqrt{\frac{1}{N} \sum_{(i,j)\in\{l\}\cup\{r\}} sm(i,j)}, \tag{10}$$

where N is the cardinality of $\mathcal{O}_l \cup \mathcal{O}_r \cup \mathcal{C}_l \cup \mathcal{C}_r$, and

$$sm(i,j) = \begin{cases} \Delta m_l(i,j)^2 & \text{if } (i,j) \in \mathcal{O}_l \cup \mathcal{S}_l, \\ \Delta m_r(i,j)^2 & \text{if } (i,j) \in \mathcal{O}_r \cup \mathcal{S}_r, \\ \frac{\Delta m_l(i,j)^2 + \Delta m_r(i\text{-}d_l,j)^2}{2} & \text{if } (i,j) \in \mathcal{R}_l, \\ \frac{\Delta m_l(i+d_r,j)^2 + \Delta m_r(i,j)^2}{2} & \text{if } (i,j) \in \mathcal{R}_r, \\ 0 & \text{if } (i,j) \in \mathcal{T}_l \cup \mathcal{T}_r, \end{cases} \tag{11}$$

where $\Delta m_v = m_v - m_{\tilde{v}}$, $v \in \{l, r\}$. If the pixel is occluded, it means that, it belongs only to one of the two views. Thus $sm(i,j)$ is computed based on one-side distortion. For the binocular suppression category, the pixel degradation in one view is masked by pixels in the other view. Only the distortion of pixels that mask their corresponding pixels in the other view will be considered. When binocular rivalry occurs, observers can see two different intensities in shutter way of the same 3D point. Hence, the average of the pixel distortions in the two views is calculated.

5. EXPERIMENTAL RESULTS

The performance of the proposed metric is evaluated using the scored LIVE 3D IQA database (phase I) [18]. The symmetric distortions that were simulated are compression using the JPEG and JP2K compression encoders, additive white Gaussian noise (WN), Gaussian blur and a fast-fading (FF) modeled based on the Rayleigh fading channel. Twenty reference stereo pairs were distorted symmetrically using the five different distortions mentioned above. With a total of 365 distorted stereo pairs (80 for each distortion except for blur we used only 45 pairs). For the WN distortion, the noise amplitude is estimated using the difference between the original and the degraded images. For the four other distortions and the original views; we assume that images are noise free. The subjective quality evaluation of the stereo image pairs in the database is followed by an objective evaluation using several state-of-the-art SIQA metrics [5–8, 10, 12, 13]. The DMOS vs the objective scores obtained by the proposed algorithm are shown in Fig. 1.

Three measures are used to evaluate the performance of the proposed metric. The Spearman's rank ordered correlation coefficient (SROCC), the linear (Pearson's) rank ordered correlation coefficient (LCC) and the Root-Mean-Squared Error (RMSE) are computed between DMOS and algorithm score DMOS_o after nonlinear regression. The non-linearity chosen for regression was a five-parameter logistic function defined in [19] as follows:

$$\text{DMOS}_o(x) = \beta_1\left(\frac{1}{2} - \frac{1}{1 + e^{\beta_2(x-\beta_3)}}\right) + \beta_4 x + \beta_5. \tag{12}$$

Each view in the database is converted into the YC_bC_r color space, which is widely used in image and video compression and processing. Furthermore, since the HVS is more sensitive to the luminance variations, only the Y component is used in the proposed metric. We fixed s to 1, ω to 15 and the graph-cut model [20] is used to estimate the disparity maps. It is worth noting that the disparity estimation model and the fixed s and ω values do not give optimal

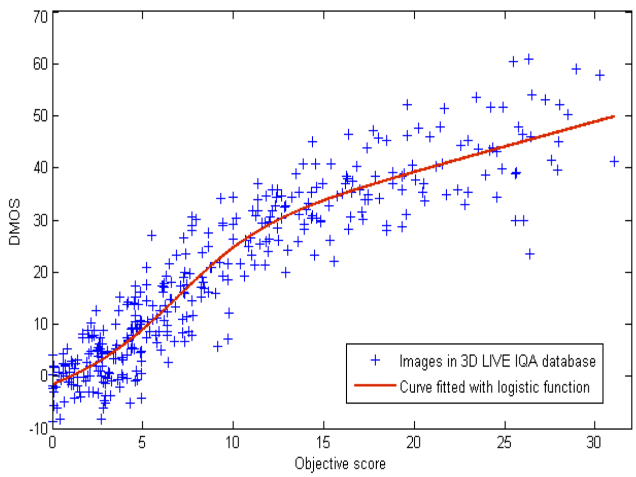

Fig. 1: DMOS *vs* objective score.

Algorithm	JP2K	JPEG	WN	Blur	FF	All
Our metric	**0.9170**	0.6595	**0.9408**	**0.9503**	**0.7745**	**0.9251**
Benoit [5]	0.9103	0.6028	0.9292	0.9308	0.6989	0.8992
Hewage [13]	0.8558	0.5001	0.8963	0.6900	0.5447	0.8140
You [7]	0.8598	0.4388	0.9395	0.8822	0.5883	0.8789
Gorley [8]	0.4203	0.0152	0.7408	0.7498	0.3663	0.1419
Shen [10]	0.2133	0.2440	0.8917	0.6586	0.2665	0.0679
Yang [11]	0.1501	0.1328	0.8471	0.3266	0.1426	0.0785
Zhu [12]	0.7708	0.2929	0.4651	0.7935	0.4752	0.6388
Akhter [6]	0.8657	**0.6754**	0.9137	0.5549	0.6393	0.3827

Table 1: Spearman's Rank Ordered Correlation Coefficient (SROCC).

performance results. Therefore, there is still room for improvement over the obtained results by optimizing the s, ω and disparity estimation.

Tables 1, 2 and 3 correspond to the performance of the proposed SIQA algorithm evaluated on the 3D LIVE IQA database. These results show that the proposed metric has better performance than the state-of-art metrics, when considering the whole database. They show also the good performance of our metric for the considered distortions. Our metric outperforms all the others considered except Akhter's metric for JPEG distortion, because the latter is specifically designed for this kind of distortion.

According to the study done in [18], 2D IQA algorithms correlate well with subjective scores when used for stereo image pairs evaluation. 3D algorithms, including our metric, do not improve the performance however. This observation is based on the nature of the 3D LIVE IQA database that contains only symmetric distortions. 2D IQA algorithms may fail in the case of asymmetric distortion when coding artefacts are different in the two decoded views. Semantically, if we consider the following case of asymmetric distortion, where the left view is kept unchanged and the right one is blurred. The 2D algorithms predict the quality by averaging the two obtained scores for the two views, hence, the predicted scores of perceived distortion depend only on the distorted level of the right image, and increase as its level of degradation increases. On the contrary, the proposed 3D metric has a different behaviour and the predicted quality depends on the high quality view and on the binocular suppression criteria. Therefore, the proposed metric may give better performance over asymmetric databases.

Algorithm	JP2K	JPEG	WN	Blur	FF	All
Our metric	**0.9402**	0.6565	**0.9502**	**0.9624**	**0.8472**	**0.9341**
Benoit [5]	0.9398	0.6405	0.9253	0.9488	0.7472	0.9025
Hewage [13]	0.9043	0.5305	0.8955	0.7984	0.6698	0.8303
You [7]	0.8778	0.4874	0.9412	0.9198	0.7300	0.8814
Gorley [8]	0.4853	0.3124	0.7961	0.8527	0.3648	0.4511
Shen [10]	0.5039	0.3899	0.8988	0.6846	0.4830	0.5743
Yang [11]	0.2012	0.2738	0.8701	0.6261	0.2824	0.3909
Zhu [12]	0.8073	0.3790	0.5178	0.7770	0.5038	0.6263
Akhter [6]	0.9059	**0.7294**	0.9047	0.6177	0.6603	0.4270

Table 2: Linear Correlation Coefficient (LCC).

Algorithm	JP2K	JPEG	WN	Blur	FF	All
Our metric	**4.4108**	4.9327	**5.1861**	**3.9315**	**6.6013**	**5.8546**
Benoit [5]	4.4266	5.0220	6.3076	4.5714	8.2578	7.0617
Hewage [13]	5.5300	5.5431	7.4056	8.7480	9.2263	9.1393
You [7]	6.2066	5.7097	5.6216	5.6798	8.4923	7.7463
Gorley [8]	11.323	6.2119	10.197	7.5622	11.569	14.635
Shen [10]	12.275	6.0216	7.2939	10.554	10.882	13.547
Yang [11]	12.697	6.2894	8.2002	12.129	11.946	15.248
Zhu [12]	7.6813	6.0684	14.720	9.1270	10.736	12.782
Akhter [6]	5.4836	**4.4736**	7.0929	11.387	9.3321	14.827

Table 3: Root-mean-squared-error (RMSE).

6. CONCLUSIONS AND FUTURE WORKS

We have proposed in this paper a novel metric for objective stereo image quality assessment. The metric is designed based on HVS modeling by exploiting the CSF and multichannel decomposition using cortex filters to derive perceptual views. Additionally, the proposed metric makes use of the BJND model to compute the just noticeable distortion within each view of the reference stereo pair. The BJND model was also used to model the binocular suppression theory. The experimental results show that the proposed metric correlates better with subjective scores than the state-of-the-art methods, at the cost of increased complexity. The experiment is performed using the 3D LIVE IQA database, which contains only symmetrically distorted images. In this work the designed metric does not take into account the disparity distortion. In future works, the proposed metric will be tested using other databases, especially those containing both asymmetric and symmetric distortions. We will also, incorporate in the metric a second term to consider the disparity artefacts. Furthermore, other models will be considered for the HVS to reduce the complexity of the algorithm.

7. REFERENCES

[1] W. Lin and C.C. Jay Kuo, "Perceptual visual quality metrics: A survey," *Journal of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.

[2] S. Daly, "The visible differences predictor: an algorithm for the assessment of image fidelity," *Digital images and human vision*, vol. 4, pp. 124–125, 1993.

[3] Y. Zhao, Z. Chen, C. Zhu, Y-P. Tan, and L. Yu, "Binocular just-noticeable-difference model for stereoscopic images," *Signal Processing Letters, IEEE*, vol. 18, no. 1, pp. 19 –22, jan. 2011.

[4] P. Campisi, P. Le Callet, and E. Marini, "Stereoscopic images quality assessment," in *EUSIPCO*, Poznan, Poland, September 2007.

[5] A. Benoit, P. Le Callet, and P. and Campisi, "Quality assessment of stereoscopic images," *EURASIP journal on image and video processing*, vol. 2008, 2009.

[6] R. Akhter, Z.M.P. Sazzad, Y. Horita, and J. Baltes, "No-reference stereoscopic image quality assessment," in *IS&T/SPIE Electronic Imaging*, 2010, vol. 7524.

[7] J. You, L. Xing, A. Perkis, and X. Wang, "Perceptual quality assessment for stereoscopic images based on 2d image quality metrics and disparity analysis," in *VPQM*, Scottsdale, Arizona, USA, 2010.

[8] P. Gorley and N. Holliman, "Stereoscopic image quality metrics and compression," in *IS&T/SPIE Electronic Imaging*, 2008, vol. 6803, p. 680305.

[9] Z.M.P. Sazzad, S. Yamanaka, Y. Kawayokeita, and Y. Horita, "Stereoscopic image quality prediction," in *QoMex*, San Diego, California, USA, 2009, pp. 180–185.

[10] L. Shen, J. Yang, and Z. Zhang, "Stereo picture quality estimation based on a multiple channel hvs model," in *CISP*, Tianjin, China, 2009, pp. 1–4.

[11] J. Yang, C. Hou, Y. Zhou, Z. Zhang, and J. Guo, "Objective quality assessment method of stereo images," in *3DTV-CON*, Potsdam, Germany, 2009, pp. 1–4.

[12] Z. Zhu and Y. Wang, "Perceptual distortion metric for stereo video quality evaluation," *WSEAS*, vol. 5, no. 7, pp. 70, 2009.

[13] C.T.E.R. Hewage and M.G. Martini, "Reduced-reference quality metric for 3d depth map transmission," in *3DTV-CON*, Tampere, Finland, 2010, pp. 1–4.

[14] A. Maalouf and M.C. Larabi, "Cyclop: A stereo color image quality assessment metric," in *ICASSP*, Prague, Czech Republic, 2011, pp. 1161–1164.

[15] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A comparative study of energy minimization methods for markov random fields," Graz, Austria, 2006, pp. 16–29.

[16] D. Scharstein and R Szeliski, "Middleburry stereo vision website," *http://vision.middlebury.edu/stereo.*

[17] W. Hachicha, A. Beghdadi, and F.A. Cheikh, "Combining depth information and local edge detection for stereo image enhancement," in *EUSIPCO*, Bucharest, Romania, 2012, pp. 250–254.

[18] A.K. Moorthy, C.C. Su, A. Mittal, and A.C. Bovik, "Subjective evaluation of stereoscopic image quality," *Signal Processing: Image Communication*, 2012.

[19] H.R. Sheikh, M.F. Sabir, and A.C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3440 –3451, 2006.

[20] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *PAMI*, vol. 23, no. 11, pp. 1222–1239, 2001.