Virtual Reality Based Just Noticeable Difference Model for Stereoscopic Images

Di Liu and Zhenzhong Chen

***Abstract*—**

***Index Terms*—Foveated just-noticeable-difference (FJND), VR-JND, luminance adaptation, contrast masking, disparity masking, stereoscopic images.**

I. Introduction

RECENTLY, as the release of commercial head-mounted devices (HMD) such as HTC VIVE and Sony PlayStation VR, virtual reality (VR) is enjoy a new booming development. VR provides a realistic and immersive simulation of a three-dimensional 360-degree environment, so that it can be widely applied into entertainment, healthcare, education, trainment, engineering, etc. However, large data capacity is a consequent problem of VR for the goal to provide realistic viewing experience. VR data is regarded as a kind of stereoscopic resourse. It is clearly advantageous to study the characteristics of the human visual system (HVS) to stereopsis for enhancing the quality of VR experience.

The visual sensitivity of the HVS is limited. The just noticeable difference (JND) or the difference threshold refers to the minimum visibility threshold below which the pixel level variations cannot be perceived by the HVS [1]. As a charactor of HVS, JND is widely applied into fields e.g. image/video compression, visual qualtiy assessment, etc.

In the literature, several computational models have been proposed for estimating pixel-wise JND for 2D image and video. There are many factors related to the visibility of threshold of a particular stimuli. Chou *et al.* proposed one of the earliest JND model in [1], which was built with background lumilance adaptation and contrast masking. Wu *et al.* proposed that the HVS is insensitive to the disorderly pattern and considered pattern complexity into JND modeling [2–4]. Visual acuity is another factor which influences JND. The fovea which is at the center of the retina has the highest density of cone and ganglion cells and the density of sensor cells drops as the retinal eccenticity increases. Thus, a foveated model that the salient part be coded with higher quality is considered in image and video compression [5, 6] for the fovea has the highest visual acuity. A specific model named foveated JND (FJND) which maps the function between JND with lumilance adaptation and fovea eccentricity was proposed by Chen *et al.* in [7]. Li *et al.* improved the FJND model by modeling the function between JND with contrast masking and fovea eccentricity in [8].

D. Liu and Z. Chen are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei 430079 P. R. China (e-mail: dliu@whu.edu.cn; zzchen@whu.edu.cn).

3D image provides an additional experience of depth to its viewers. Depth perception is the result of several depth cues, including monocular ones, e.g. occlusion, perspective convergence, motion parallax, texture gradient, etc. and binocular ones e.g. convergence and binocualr disparity [9]. Among these cues, binocular disparity, which refers to the lateral displacement of an object seen by the left and right eyes, is widely used in the stereopsis to provide viewers with 3D experience. Thus, binocular disparity is the mostly used as the representative depth feature for stereoscopic viewing.

Some studies have been done to investigate the characteristic of HVS to depth feature. Eye fixation experiments in [10–12] all indicated a depth bias that objects closer to the viewer attract attention earlier than distant objects and always attract the most eye fixation. Yand *et al.* found that the distributions of natural disparities in both indoor and outdoor scenes are centered at zero, have high peaks, and span about 5 deg [13]. In [14, 15], Silva *et al.* investigated the sensitivity of the HVS for depth cues in 3D dispalys and proposed the just noticeable depth difference (JNDD) model for suppressing the unnecessary spatial depth details of depth maps. The JNDD modle is adopted by Jung *et al.* for depth sensation enhancement in [16, 17]. In [18], Zhao *et al.* considered binocular combination and rivalry and proposed the a binocular JND (BJND) model in response to asymmetric noises in a pair of stereoscopic images.

To our best knowledge, none of the existing JND models are developed for stereoscopic images which mapping the JND of pixel to the corresponding relative disparity between the pixel and the eye fixation. In this paper, we experimently investigate the function of JND to depth by an HTC VIVE and propose an improved and extended FJND model for stereoscopic image which take relative disparity into the JND function. As it is the first research to investigate JND in virtual reality, the proposed model is named virtual reality based just-noticeable-distortion (VR-JND) model.

The rest of this paper is organized as follows. In Section II, a brief introduction to the 2D just-noticeable-distortion models considered in this paper is provided. Mathematical model of the proposed foveated disparity just-noticeable-distortion (VR-JND) characteristic is presented in Section III. Section IV presents the experimental validation of the derived mathematical models and Section V concludes the paper.

II. FJND Model

It is widely accepted that background luminance adaptation and spatial contrast masking are two mainly factors which affect JND [1, 3]. Since the visual acuity decreases when the distance from the fovea increases, the visibility threshold increases with increased eccentricity. JND that varies according to the eccentricity of the pixel to the fovea on retina is named foveated JND (FJND). According to the works in [7, 8], the FJND is defined as

$$FJND = F_1 + F_2 - C^{gr} \cdot \min\{F_1, F_2\} \tag{1}$$

where C^{gr} is gain reduction parameter that measures the overlapping effect and is set as 0.8 empirically, F_1 and F_2 represent foveated luminance adaptation effect and foveated contrast masking effect respectively as

$$F_1 = f_1 \cdot m_1 \tag{2}$$

$$F_2 = f_2 \cdot m_2 \tag{3}$$

$$f_1(bg) = \begin{cases} 14 \times \left(1 - \left(\frac{bg}{127}\right)^{1/2}\right) + 2 & \text{if } bg \leq 127 \\ \frac{3}{128} \times (bg - 127) + 2 & \text{if } bg > 127 \end{cases} \tag{4}$$

$$f_2(bg, mg) = mg \times \alpha(bg) + \beta(bg) \tag{5}$$

$$m_1(bg, v, e) = (2 - S_f(v, e))^{\eta_1(bg)} \tag{6}$$

$$m_2(eh, v, e) = \left(\frac{1 + 0.05}{S_f(v, e) + 0.05}\right)^{\eta_2(eh)} \tag{7}$$

$$S_f(v, e) = \frac{f_m(v, e)}{f_m(v, 0)} \tag{8}$$

$$f_m(v, e) = \min\left(\frac{2.3 \times \ln(64)}{0.106 \times (e + 2.3)}, \frac{\pi v}{360}\right) \tag{9}$$

$$\eta_1(bg) = 0.5 + \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\log_2(bg + 1) - \mu)^2}{2\sigma^2}\right) \tag{10}$$

$$\eta_2(eh) = \left(\log\left(\frac{eh}{255} + 1.5\right)\right)^2 + 0.66 \tag{11}$$

III. VR-JND Model

A. Experimental Setup

Instead of displaying images on monitors as in [18], we use HTC VIVE glasses to present stereoscopic images. Bacause we aim to study the influence of relative disparity on JND in our experiment, and VR HMD has the advantage to fix the viewing distance which leads to a more reliable JND model. HTC VIVE has an OLED with resolution of 1080*1280 per eye and the refreshing rate of 90 Hz. By showing two images on two seperate screen before eyes, VR HMD can provide an imersive 3D environment. The distance from the eyes of viewers and the screens is fixed, when viewers fixed their attention to a point in the virtual 3D environment, the relative disparity between this fixation and other point is easy to calculate and always even. The relationship between the point in space and points in two VR screens is shown as Fig. 1.

Our experiments are divided into two parts. In the first part, we investigate the effect of relative depth on luminance adaptation, which is represented by d_1. In the second part, the relationship between relative depth and contrast masking effect is investigated and represented by d_2.

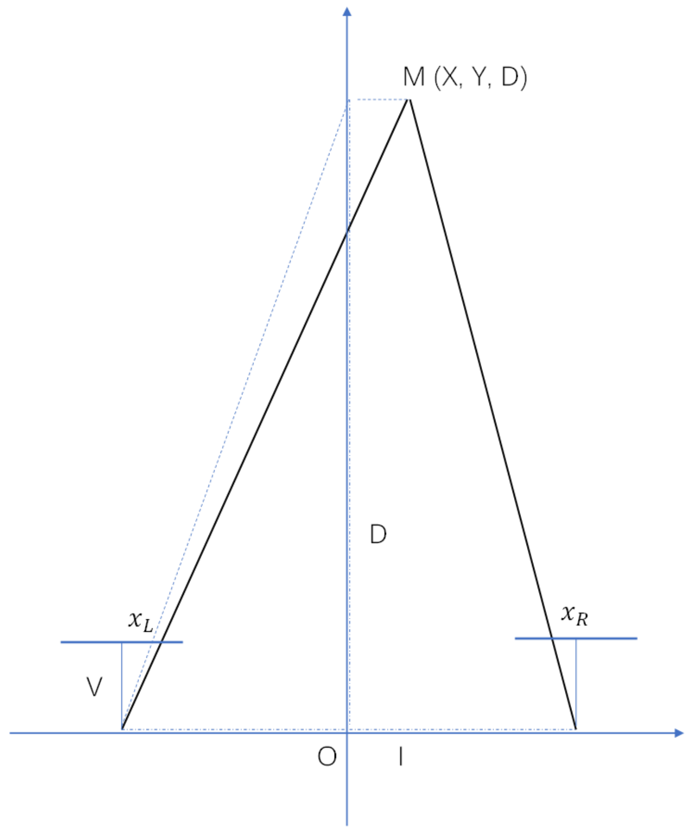

Fig. 1. Stereopsis of VR HMD.

1) Luminance Adaptation: As is known that the background luminance bg is a factor which affect the visual JND, thus it is also investigated in the VR JND experiments. Based on the aforementioned background information, we design an image pattern to explore the correlation of background luminance bg, eccentricity e, relative disparity d and JND T. The experiment setup is shown as Fig. 2, and the sketch map of the VR stimuli is shown as Fig. 3. There is a red cross at the center of the field of view at a constant viewing depth of D. During the whole experiment, subjects are requested to fix their attention at this red cross. During each trial, a chessboard noise will appears nearby the red cross in the virtual 3D space. Further more, in order to avoid the influence of predefined location of noise on subjects and help subjects keep their fixation at the center of red cross, the location of the noise is random and depends on three elements: the relative depth d to the red cross in the viewing direction, the retinal eccentricity e between the noise and the fixated red cross, and the central angle of noise to the center. As is known that stereopsis is generated by fusion of similar texture and caused by the dispartiy of the same pixel from different views. However, pure gray background is insufficient to make a three-dimensional experimence for the chessboard noise. To solve this problem, we put a square at center of the field and at the same depth field of the red cross as shown in Fig. 3 to help construct stereopsis. As mentioned, the visual acuity is identical at arbitrary angle of the circle of the same visual eccentricity. Thus JND in the luminance adaptation experiement is determined by the

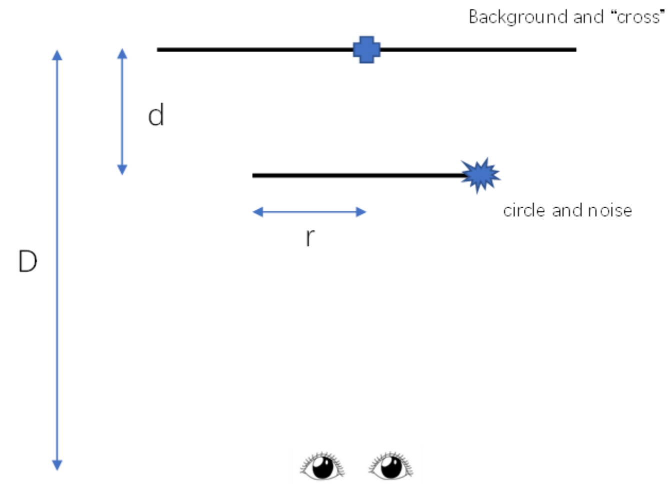

Fig. 2. Sketch of luminance adaptation experiment.

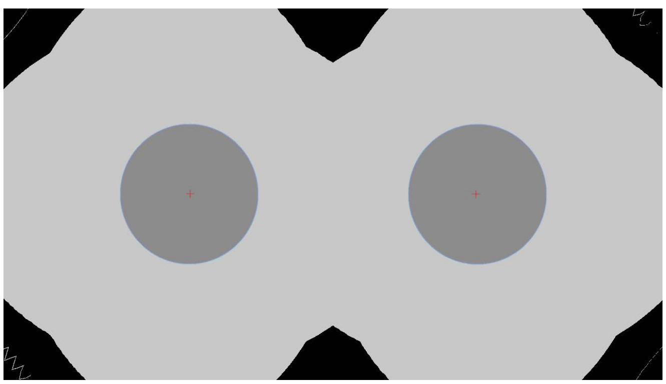

Fig. 3. Stimuli in the VR glasses of luminance adaptation experiment.

background luminance bg, eccentricity e and relative depth d.

2) Contrast Masking: To investigate the relation The pattern consists of a circular area whose radius is calculated with viewing distance and eccentricity,.

IV. Experimental Results

V. Conclusion

References

[1] C. H. Chou and Y. C. Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 6, pp. 467–476, 1995.

[2] J. Wu, W. Lin, G. Shi, X. Wang, and F. Li, "Pattern masking estimation in image with structural uncertainty," *IEEE Trans. on Image Process.*, vol. 22, no. 12, pp. 4892–4904, 2013.

[3] J. Wu, G. Shi, W. Lin, and A. Liu, "Just noticeable difference estimation for images with free-energy principle," *IEEE Trans. Multimedia*, vol. 15, no. 7, pp. 1705–1710, 2013.

[4] J. Wu, L. Li, W. Dong, G. Shi, W. Lin, and C. C. J. Kuo, "Enhanced just noticeable difference model for images with pattern complexity," *IEEE Trans. Image Process.*, vol. PP, no. 99, pp. 1–1, 2017.

[5] Z. Wang and A. C. Bovik, "Embedded foveation image coding," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1397–1410, 2001.

[6] Z. Wang, L. Lu, and A. C. Bovik, "Foveation scalable video coding with automatic fixation selection," *IEEE Trans. Image Process.*, vol. 12, no. 2, pp. 243–54, 2003.

[7] Z. Chen and C. Guillemot, "Perceptually-Friendly H.264/AVC Video Coding Based on Foveated Just-Noticeable-Distortion Model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 806–819, 2010.

[8] Y. Li, H. Liu, and Z. Chen, "Perceptually-lossless image coding based on foveated-jnd and h.265/hevc," *J. Visual Commun. Image Represent.*, vol. 40, pp. 600–610, 2016.

[9] Blake and Randolph., *Perception*. McGraw-Hill, 1990.

[10] L. Jansen, S. Onat, and P. König, "Influence of disparity on fixation and saccades in free viewing of natural scenes," *J. Vis.*, vol. 9, no. 1, pp. 29–29, 2009.

[11] C. Lang, T. Nguyen, H. Katti, K. Yadati, M. Kankanhalli, and S. Yan, "Depth matters: Influence of depth cues on visual saliency," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2012, pp. 101–115.

[12] J. Wang, P. L. Callet, S. Tourancheau, V. Ricordel, M. Perreira, and S. Da, "Study of depth bias of observers in free viewing of still stereoscopic synthetic stimuli," *J. Eye Movement Res.*, vol. 5, no. 5, pp. 10–13, 2012.

[13] L. Yang, A. C. Bovik, and L. K. Cormack, "Disparity statistics in natural scenes," *J. Vis.*, vol. 8, no. 11, pp. 1–14, 2008.

[14] D. V. S. X. De Silva, E. Ekmekcioglu, W. A. C. Fernando, and S. T. Worrall, "Display dependent preprocessing of depth maps based on just noticeable depth difference modeling," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 2, pp. 335–351, 2011.

[15] D. V. S. X. De Silva, W. A. C. Fernando, G. Nur, and E. Ekmekcioglu, "3d video assessment with just noticeable difference in depth evaluation," in *Proc. IEEE Int. Conf. Image Process.*, pp. 4013–4016.

[16] S. W. Jung, "A modified model of the just noticeable depth difference and its application to depth sensation enhancement," *IEEE Trans. Image Process.*, vol. 22, no. 10, p. 3892, 2013.

[17] S. W. Jung and S. J. Ko, "Depth sensation enhancement using the just noticeable depth difference," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3624–3637, 2012.

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