Virtual Reality Based Just-Noticeable-Distortion Model for Stereoscopic Images

Di Liu and Zhenzhong Chen

***Abstract*—Depth is important information in real stereoscopic viewing. As the development of stereoscopic resources and virtual reality technology, the role of depth information in the human visual system is in need to be investigated for the requirement to make stereoscopic data perceptually friendly. In this paper, we develop a vr based jnd model to quantify the perceptual redundancy of image in the vr display environment. The proposed model considers both luminance adaptation and contrast masking. Furthermore, the effects of both foveation and depth on luminance adaptation and contrast masking are investigated and well fitted.**

***Index Terms*—Foveated just-noticeable-distortion (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 distortion (JND) or the distortion 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].

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 model 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-

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).

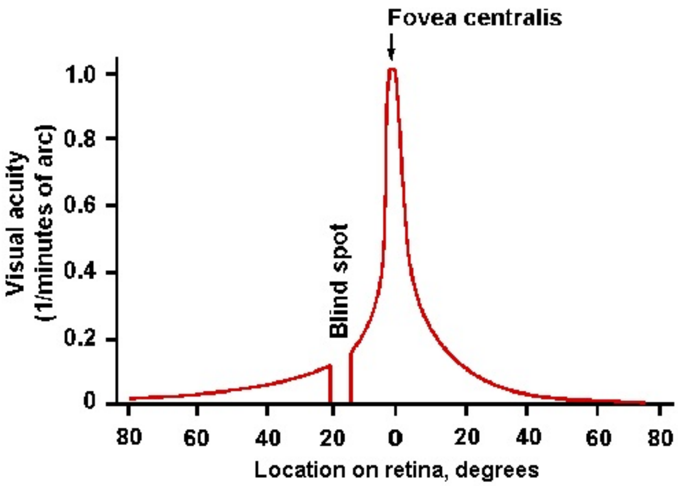


Fig. 1. Visual acuity as a function of retinal eccentricity.

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 main factors which affect JND. On this basis, the first foundation work of SJND was done in [1]. However, retinal eccentriciy is another factor which affects JND. The density of the cone and ganglion cells drops with the increased retinal eccentricity. As a result, the visual acuity decreases when the distance from the fovea increases as shown in Fig. 1 and the visibility threshold increases with increased eccentricity. Given the fixation point (x_f, y_f) and viewing distance v, the retinal eccentricity of a certain point $P\left(x, y\right)$ can be calculated as:

$$e\left(v, p\right) = \tan^{-1}\left(\frac{d}{v}\right) \tag{1}$$

where

$$d = \sqrt{\left(x - x_f\right)^2 + \left(y - y_f\right)^2} \tag{2}$$

The JND that varies according to the eccentricity of the pixel to the fovea on retina is named foveated JND (FJND). The usual method to predict eye fixtion for images is by using saliency detection model, like Itti's model in [19]. Saliency is the result of attentional selection that intermediate and higher visual processes appear to select a subset of the available sensory information before further processing. Human alternate saccades and visual fixation to focus the eyesight on objects of interest, the fixation is projected to the centre of fovea which has the highest visual acuity. Thus, it offers human the ability to interpret complex scenes in real time, despite the limited speed of the neuronal hardware available for such tasks. JND is used to remove perceptual redundancy in image compression, exploiting the foveation property that increase the visual distortion threshold for parts of image which are unlikely to be fixated is perceptually friendly. According to the works in [7, 8], the FJND is defined as

$$FJND = F_1 + F_2 - C^{gr} \cdot \min\{F_1, F_2\} \tag{3}$$

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

$$F_1 = f_1 \cdot m_1 \tag{4}$$

$$F_2 = f_2 \cdot m_2 \tag{5}$$

where m_1 and m_2 measure the luminance adaptation and spatial contrast masking effect respectively, f_1 and f_2 are the foveation models, which calculate the foveation effect of m_1 and m_2 respectively.

III. VR-JND Model

Depth is another essential element in the real world viewing. In [20], it was demonstrated that human eyes are less sensitive to the depth difference when the position of the displayed image become closer to or farther away from the display screen and an approximated JNDD formula was provided. The JNDD model can be used for depth sensation enhancement [14]. As demonstrated in by eye movement experiments in [], depth is another important feature beside color that affect visual attention in the stereoscopic view. It is named the depth-bias that human tend to pay attention to objects closer to them. Inspired by these work, we propose that depth might also have an effect on the JND in the 3D case. In this section, we do experiments with VR equipment to investigate the depth effect on JND modeling.

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. The 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. 2.

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.

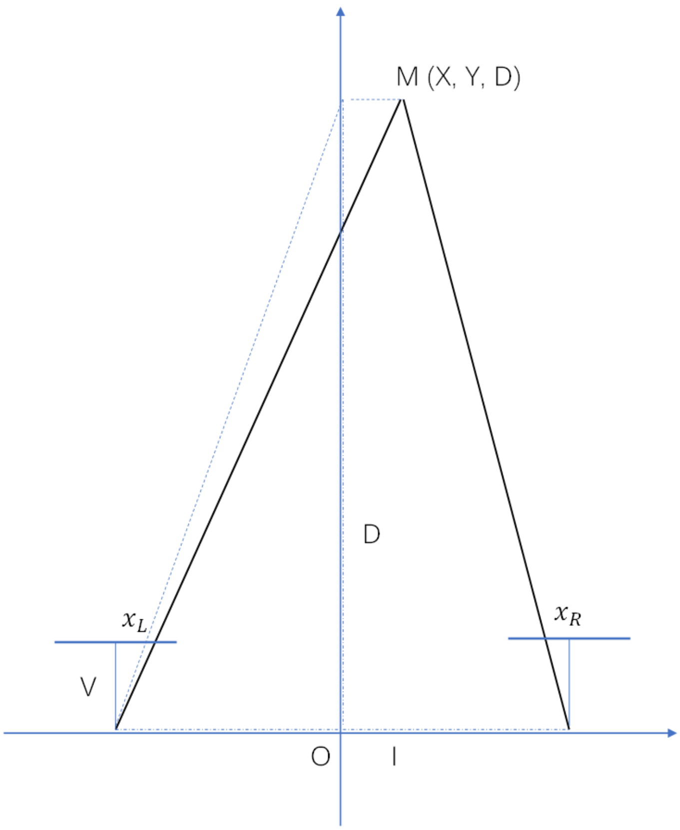


Fig. 2. Stereopsis of VR HMD.

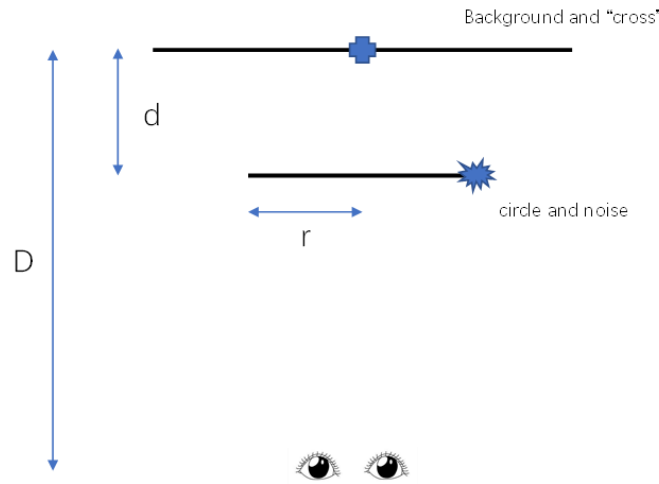


Fig. 3. Sketch of luminance adaptation experiment.

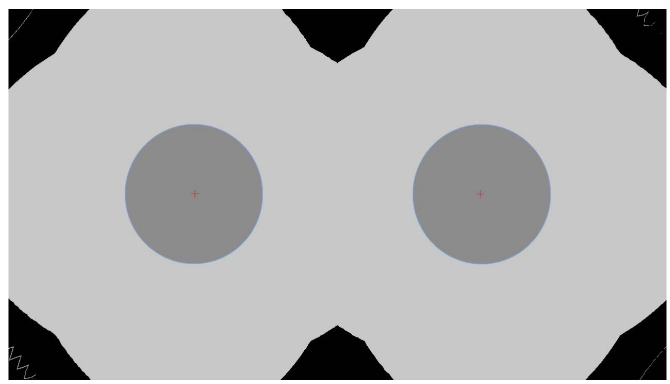


Fig. 4. Stimuli in the VR glasses of luminance adaptation experiment.

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 depth d and JND T. The experiment setup is shown as Fig. 3, and the sketch map of the VR stimuli is shown as Fig. 5. 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 noise as shown in Fig. 4 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 background luminance bg, eccentricity e and relative depth d.

During each trial, bg, e and d of the stimulus are randomly selected. At the beggining of the test, the noise amplitude was set as 0, which means the noise area was invisible to subjects. With their attention fixed at the red cross, subjects were asked to increase the noise amplitude by 1 each time with the keyboard untill the noise area became visible and the corresponding amplitude was recorded as the just noticeable threshold A for the current stimulus. 8 subjects with normal vision took part in this experiments. Each subject was asked to do 80 trials and they were informed that they can take a rest duirng the whole experiment once their eyes were fatigued.

2) Contrast Masking: To investigate the effect of depth on contrast masking, we designed the experimental stimuli as Fig. 5. The basic setup of contrast masking experiments are similar with the luminance adaptation experiments, the difference is that in the contrast masking, the pattern consists of a circular area instead of the square. The red cross is at the viewing distance D, the distance of the circular area to the red cross is the relative depth d and the radius of the cirlular area can be calculated as

$$r = (D - d) \cdot \tan(e) \tag{6}$$

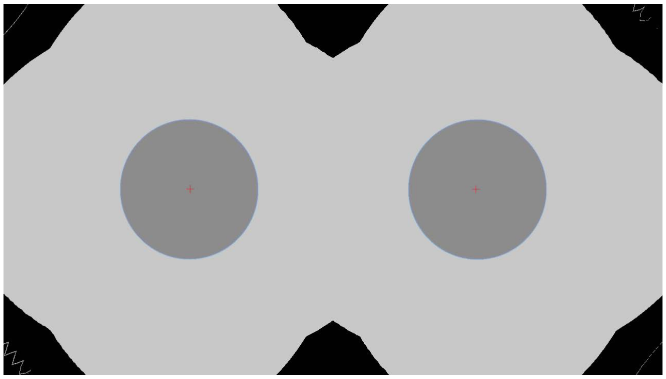


Fig. 5. Stimuli in the VR glasses of luminance adaptation experiment.

The chessboard noise will be located along the edge of the circle and at the same depth of the circle. Luminance contrast between the circle and the background is eh. The circle can be used to construct stereopsis. In the contrast masking experiments, JND is determined by the luminance contrast eh, eccentricity e and relative depth d. Testing process is similar with that in the luminance adaptation experiments, 8 subjects with normal vision took part in this experiment. Each subject was asked to do 80 trials and they were informed that they can take a rest duirng the whole experiment once their eyes were fatigued.

IV. Experimental Results and Analysis

A. Luminance Adaptation

Subjective experiment result of luminance adaptation part is shown in Fig. 6. The luminance adaptation effect is influenced by three factors, the relationship can not be illustrated geometrically. However, we can illustrate the relationship between jnd and two factors with a 3D surface by averaging on the rest dimension. As shown in Fig. 6, the surface is averaging on the depth dimension, and it shows the relationship between jnd with background luminance and retinal eccentricity. As luminance increases, the JND drops at the beginning and increases when the luminance gets higher. The extreme point is at around the luminance of 50. In order to form the function between jnd and three factors, firtly, we investigate the relationshio between jnd and one fator once as shown in the Fig. 7. The mean bg-jnd of y axis represents the average jnd value for the variable of x axis. The fitting curve of jnd to background luminance, retinal eccentricity and depth can be represented by m_1, f_1 and d_1 respectively as

$$m_1 = a_1 \cdot bg^2 + a_2 \cdot bg + a_3 \tag{7}$$

$$f_1 = a_4 \cdot e^{a_5} + a_6 \tag{8}$$

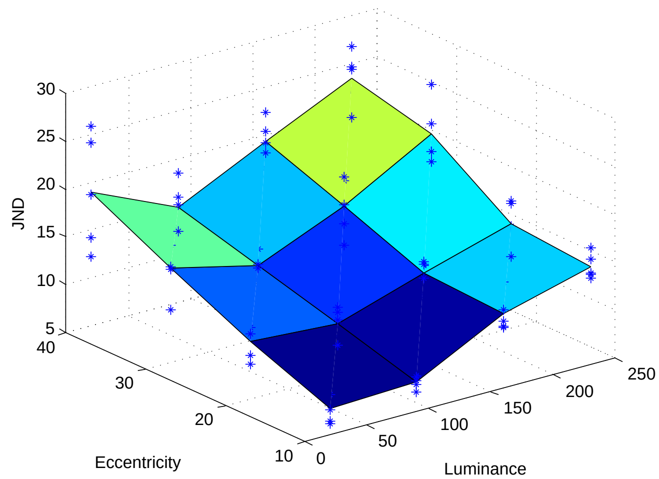
$$d_1 = a_7 \cdot d + a_8 \tag{9}$$

with $a_1 = 0.000269$, $a_2 = -0.03652$, $a_3 = 13.5$, $a_4 = 0.037$, $a_5 = -1.5$, $a_6 = 9.65$, $a_7 = 0.0466$, $a_8 = 0.0216$. The fitting precision for luminance adaptation is list in Tab. I. and the luminance adaptation effect can be written by

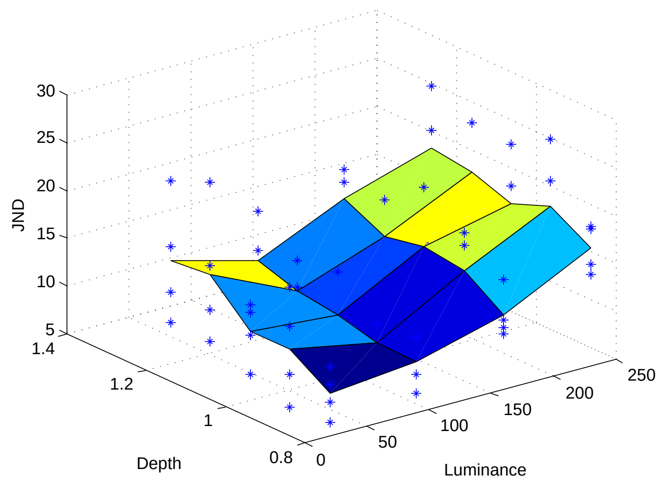
$$F_{bg} = m_1 \cdot f_1 \cdot d_1 \tag{10}$$

TABLE I
Fitting Precision of Luminance Adaptation Effect.

Function	RMSE	R-square	Adjusted R-square
bg-jnd to bg	1.317	0.946	0.837
bg-jnd to e	0.623	0.991	0.972
bg-jnd to d	0.330	0.967	0.956
bg-jnd to bg, e and d	2.238	0.812	0.807



(a)



(b)

Fig. 6. Subjective experiment result of luminance adaptation part.

B. Contrast Masking

Subjective experiment result of contrast masking part is shown in Fig. 8. The contrast masking effect is also influenced by three factors, the relationship can not be illustrated geometrically. However, we can illustrate the relationship between jnd and two factors with a 3D surface by averaging on the rest dimension. As shown in Fig. 8, the surface is averaging on the depth dimension, and it shows the relationship between jnd with luminance contrast and retinal eccentricity. As contrast increases, the JND increases as the luminance gets higher. In order to form the function between jnd and three factors, firtly, we investigate the relationshio between jnd and one fator once as shown in the Fig. 9. The mean eh-jnd of y axis represents the average jnd value for the variable of x axis. The fitting

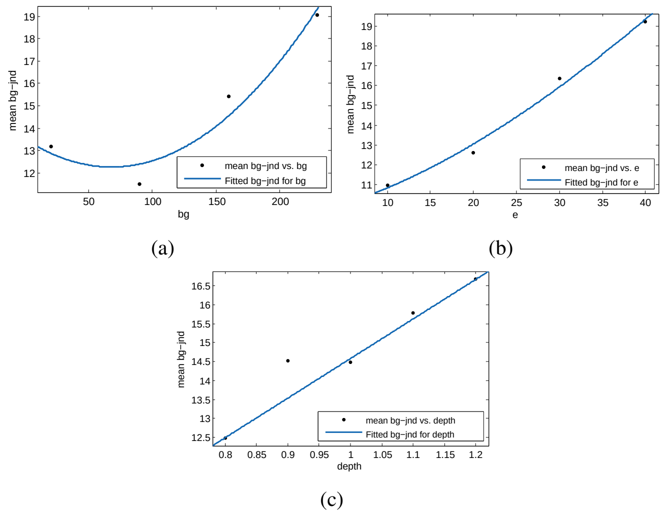


Fig. 7. Fitting curves for luminance adaptatation experimental result.

curve of jnd to contrast, retinal eccentricity and depth can be represented by m_2, f_2 and d_2 respectively as

$$m_2 = b_1 \cdot eh + b_2 \tag{11}$$

$$f_2 = b_3 \cdot e + b_4 \tag{12}$$

$$d_2 = b_5 \cdot d + b_6 \tag{13}$$

with $b_1 = 0.0525$, $b_2 = 14$, $b_3 = 0.08$, $b_4 = 13.3$, $b_5 = 0.0798$ and $b_6 = -0.0158$. The fitting precision for contrast masking is list in Tab. II. The contrast masking effect can be written by

$$F_{eh} = m_2 \cdot f_2 \cdot d_2 \tag{14}$$

TABLE II
FITTING PRECISION OF CONTRAST MASKING EFFECT.

Function	RMSE	R-square	Adjusted R-square
eh-jnd to eh	0.556	0.923	0.885
eh-jnd to e	0.237	0.969	0.954
eh-jnd to d	0.304	0.992	0.990
eh-jnd to eh, e and d	3.180	0.485	0.472

C. VR-JND

Taking both the luminance adaptation effect and contrast masking effect into consideration, the final VR-JND can be refined as

$$VR-JND = F_{bg} + F_{eh} - C^{gr} \cdot \min\{F_{bg}, F_{eh}\} \tag{15}$$

where F_{bg} represents the luminance adaptation effect, C^{gr} is gain reduction parameter that measures the overlapping effect, and is set as 0.8 empirically.

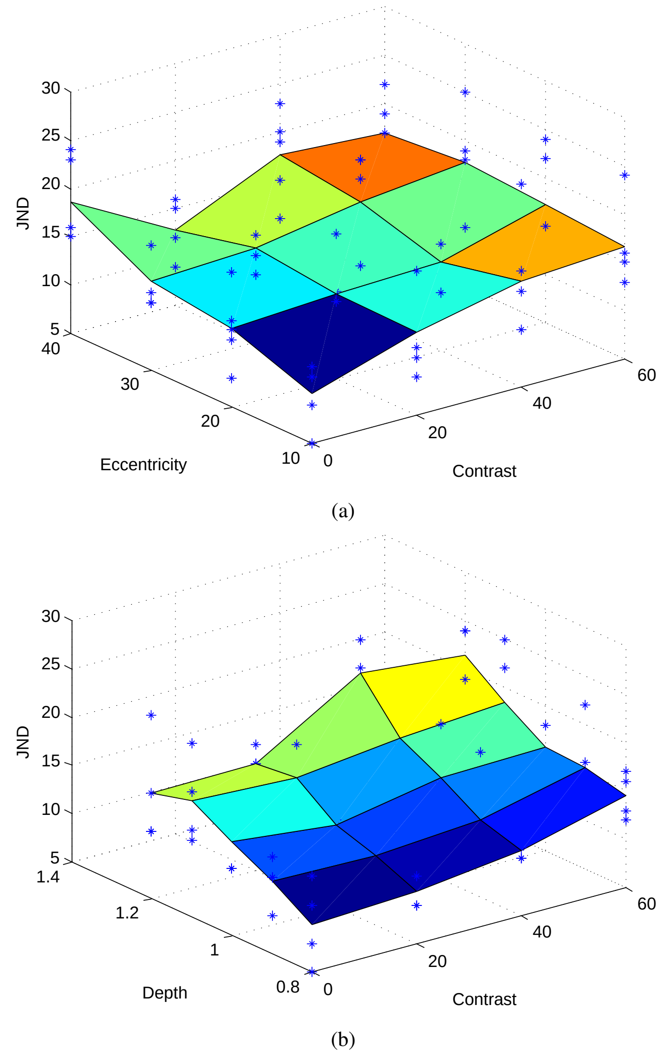


Fig. 8. Subjective experiment result of contrast masking part.

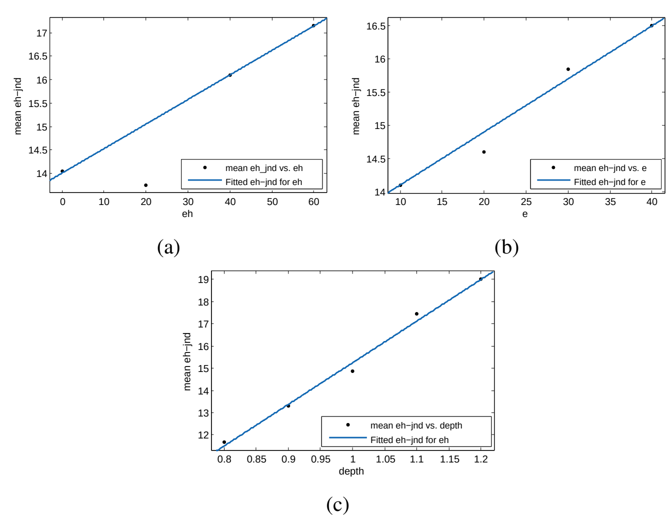


Fig. 9. Fitting curves for contrast masking experimental result.

V. CONCLUSION

In this paper, we propose a virtual reality based just-noticeable-distortion model for stereoscopic images based on psychophysical experiments. The foveation effect and depth effect for luminance adaptation and contrast masking are both model in the VR viewing condition. It is the first time that the JND porperty of HVS is investigated in the VR environment. We got a good fitting precision for both the luminance adaptation and contrast masking effects.

.

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.

[19] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 11, pp. 1254–1259, 1998.

[20] V. D. Silva, A. Fernando, S. Worrall, and H. K. Arachchi, "Sensitivity analysis of the human visual system for depth cues in stereoscopic 3-d displays," *IEEE Trans. Multimedia*, vol. 13, no. 3, pp. 498–506, 2011.