Visual Discomfort Estimation for Stereoscopic 3D Contents

Chung-Te Li, Chien Wu, Cheng-Yuan Ko, Chen-Han Chung, and Liang-Gee Chen*

DSP/IC LAB, Graduate Institute of Electronic Engineering, National Taiwan University

***Abstract*—Stereoscopic 3D displays are widely applied for home entertainments. However, some viewers complain uncomfortable 3D perception when they watch 3D contents. For generating comfort stereoscopic 3D contents, we proposed three qualitative models for measuring the discomfort. Three experiments for 1) limitation of the viewing angle 2) variance of the viewing angle for the whole scene 3) variance of the viewing angle inside an object are designed for generating the models in this paper. On the basis of the experiments, we find three simple features in terms of thresholds for estimating the potential visual discomfort for a given stereoscopic 3D content. We believe these thresholds can be applied for the production of comfort 3D contents, especially for automatic or semi-automatic generation of 3D contents.**

I. INTRODUCTION

In the near future, 3D displays will become more popular in our daily life. A discussion for its perceptual experience is also highly demanded now. Depth map generation from 2D images or videos (2D-to-3D) is still a very hot topic now. Researchers have paid their attentions on depth generation from different depth cues. Various depth generation techniques have been proposed in the previous works [1]-[5]. These works focus on some pre-determinate depth cues and get good results of the depth information when the pre-determinate depth cue is significant in the given image or video. Some time-consuming semi-automatic generation of depth information are also proposed to provide more accurate depth information [6]-[8]. However, there are still some viewers complain 3D discomforts. To analyze how the discomfort occurs and suppress it is emergent. Nojiri et al. [9] proposed a fatigue model on the basis of the spatial-temporal distribution of disparity. Sohn et al. [10] proposed another model on the basis of visual attention and the amount of disparity. Richardt et al. also propose a model [11] on the basis of the binocular coherence. These models estimate different kinds of visual discomfort very well. However, there are still some kinds of discomfort cannot be estimated by the above methods. We try to explore another approach to estimate the discomfort. In this paper, hypotheses of three simplified geometric analysis for depth perception quality are mentioned. The details about the simplified geometric model will be shown in section II. Then, some experiments are designed in section III. The experiment result and its discussion are in section IV. Finally, we conclude and discuss future works and possible applications in the last section.

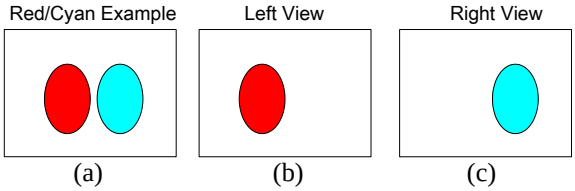

Fig. 1 Red-cyan anaglyph (a) mixed (b) left view (c) right view

II. SIMPLIFIED GEOMETRIC DEPTH PERCEPTION MODEL

We use red-cyan anaglyphs to introduce our proposed model in the following figures, just as shown in Fig. 1. In the following discussions, we will mainly discuss three geometrical features for display a 3D sequence on a stereoscopic type display.

A. Limitation of the Viewing Angle

In this paragraph, we analyze the limitation of view angle difference. The viewing angle difference is caused by the

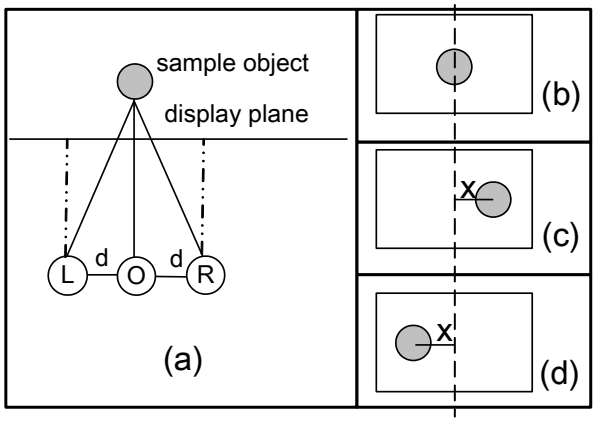

Fig. 2 setting of rendered views (a) object behind display (b) original view (c) rendered right view (d) rendered left view (L: left eye; R: right eye; O: original view)

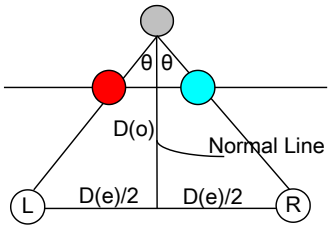

Fig. 3 The definition of viewing angle

limitation of the rendered left and right view. Both the two views need to be displayed just on the display plane. For simplicity, we assume that the rendering views are set as Fig. 2. The distance from the left view to the original view is the same as right view to original view. So, in Fig. 2 (c) and (d), the shifted distance *x* is the same.

Our proposed model decomposes all the test images into several sample objects. For each object, the we denote the geometric depth perception in terms of viewing angle as shown in Fig. 3. The viewing angle of the sample object (i.e. the gray circle in Fig. 3) is θ with reference to the normal line, where,

$$\theta = \tan^{-1}(\frac{D(e)/2}{D(o)}). \quad (1)$$

In (1), *D(o)* means the distance from sample object to viewer, and *D(e)* is the interocular distance of the viewer.

We assume that there is an upper bound of θ for comfortably viewing in 3D for near objects since viewer usually feel uncomfortable when some objects in the scene are too near. For far objects, there is no limitation except for $\theta \geq 0$. That means a sample object will make viewers uncomfortable only in the some cases in Fig. 4 (c). For objects behind or on the display, the limitation of viewing angle doesn't exist.

We'll design an experiment in section III to prove that our assumption, and then provide a threshold for the viewing angle for the visual comfort.

B. *Variance of theViewing Angle for the Whole Scene*

In this paragraph, we analyze the variance of the viewing angle range variance for the whole scene. As in part A, we also assume rendering views are set as shown in Fig. 2. The view angle range is defined in Fig. 5. We also decompose the

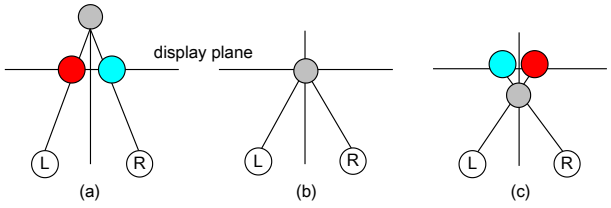

Fig. 4 (a) object behind display (b) object on display (c) object before display (L: left eye; R: right eye; gray circle: target object; Red circle: the left view on display; Cyan circle: the right view on display)

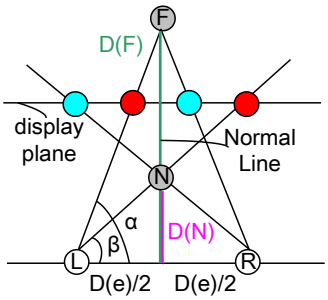

Fig. 5 The variance of the viewing angle for whole scene (L: left eye; R: right eye; F: furthest object, N: nearest object)

whole image into lots of sample objects. The nearest one, *N*, and the farthest one, *F*, are highlighted. In Fig. 5, for simplicity, we discuss all the viewing angles with reference to the line of eyes, $\overrightarrow{LR}$. Thus, for the nearest object *N*, the view angle is β. For the farthest object *F*, the view angle is α, where

$$\alpha = \tan^{-1}(\frac{D(e)/2}{D(F)}), \beta = \tan^{-1}(\frac{D(e)/2}{D(N)}). \quad (2)$$

In (2), *D(N)* means the distance from object N to $\overrightarrow{LR}$, *D(F)* means the distance from object F to $\overrightarrow{LR}$, and D(e) is also the interocular distance of the viewer. Notably, for all the viewing angles γ of other objects in the scene, the following equation is always satisfied.

$$\beta \leq \gamma \leq \alpha \quad (3)$$

Since viewers usually complain that they feel uncomfortable because the protrusion is too strong, we assume that |α-β| is also an important criterion of the view angle of the whole scene. We'll also design an experiment in section III to prove that our proposed model works, and then provide a threshold for the viewing angle for the visual comfort.

C. *Inside-object Viewing Angle Differance*

In this paragraph, we analyze the view angle difference inside a single object. Rendering views are also set as shown in Fig. 2. For simplicity, we assume that the average depth of the given object is just constant. In Fig. 6, considering the gray circle *P* is a pixel of the given object. The view angles of left and right eyes are both θ. Where

$$\theta = \tan^{-1}(\frac{D(e)/2}{D(P)}). \quad (4)$$

D(P) means the assigned disparity of pixel *P*, and *D(e)* is also the interocular distance of the viewer.

We denote pixel set *I* as the set of all the pixels inside an given object. According to our assumption, we can derive:

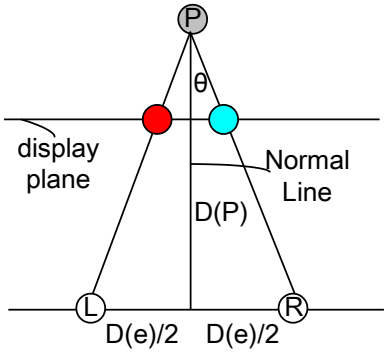

Fig. 6 The viewing angle difference for pixel P (L: left eye; R: right eye; P: sample pixel); Red circle: the left view on display; Cyan circle: the right view on display)

$$D(I) = \frac{1}{|I|}\sum_{P \in I} D(P) \text{ is constant.} \quad (5)$$

where *D(I)* means average disparity of *I*.

We define that the inside-object viewing angle distance (*IOVAD*) is proportion to the average angle θ. It is shown in the following equation, i.e.,

$$IOVAD = \frac{1}{|I|}\sum_{P \in I} \tan^{-1}\left(\frac{D(e)/2}{D(P)}\right) = \frac{1}{|I|}\sum_{P \in I} \cot^{-1}\left(\frac{D(P)}{D(e)/2}\right) \quad (4)$$

Through the concavo-convexity of $\cot^{-1}$ function, the minimum angle difference happens if and only if the following equation are satisfied:

$$D(P) = D(I) \quad \forall P \in I \quad (5)$$

III. Experiment Design

In this section, we design several experiment to support our proposed model. We choose NVIDIA 3D vision as the solution of stereoscopic display. In contrast, subjects of all the experiments are asked to be 50cm from the display plane. The size of the display is 22". We design our test pattern by placing some sample circles with difference depth and use the ease of 3D fusion as the criterion for visual discomfort.

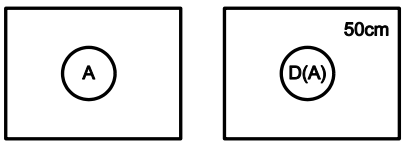

Fig.7 The sample test pattern of experiment A

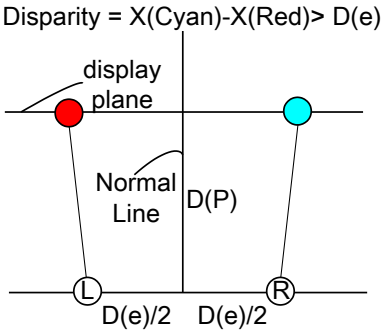

Fig. 8 Hyper-divergence pattern example

A. Limitation of the Viewing Angle

The test pattern is shown in Fig. 7. Give a sample object A, adjust the distance *D(A)*, and then ask subjects the ease fusion range of *D*.

In our experiment, we tested each subject by 10 kinds of D(A): 250m(approximated ∞), 100cm, 66.67cm, 57.13cm, 50cm, 44.45cm, 40cm, 33.33cm, 25cm, 20cm, and 12.5cm. The corresponding disparities of left and right view are -80, -40, -20, -10, 0, 10, 20, 40, 80, 120, and 240 pixels, respectively. We also tested hyper-divergence samples with -120 and -240 pixels like in Fig. 8. The disparity is larger than the interocular distance of the viewer.

B. Variance of theViewing Angle for the Whole Scene

For simplicity, our test pattern only contains two objects. The test pattern is shown in Fig. 9. Give two sample objects A and B, adjust the distance *D(A)* and *D(B)*, and then ask subjects the ease fusion range of *D(A)* and *D(B)*.

In our experiment, we tested each subject by 9x9 kinds of *D(A)* and *D(B)*: 250m, 100cm, 66.67cm, 57.13cm, 50cm, 44.45cm, 40cm, 33.33cm, and 25cm. The corresponding disparities of left and right view are -80, -40, -20, -10, 0, 10, 20, 40, and 80 pixels, respectively. The subjects are asked to focus on A or B, and they will tell if the other object can be fused.

C. Inside-object Viewing Angle Differance

For simplicity, our test pattern only contains two objects. The test pattern is shown in Fig 10. Give two sample objects A and B, fix the average distance *D(A)* and *D(B)*, increase the depth variance of *A* and then ask subjects the ease fusion range of *D(A)* with *D(B)* is well fused. Notably, *D(B)* is a fixed fuse reference, and its value is the same as average(*D(A)*).

In our experiment, we test each subject by 8 kinds of D(A): about 100cm, 66.67cm, 57.13cm, 50cm, 44.45cm, 40cm, 33.33cm, and 25cm. For each D(A), it's variance is slowly

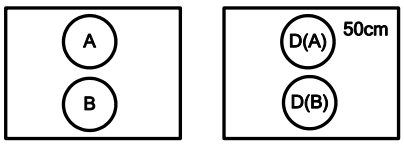

Fig. 9 The sample test pattern of experiment B

TABLE I THE RESULT OF EXPERIMENT B

DISP	*-80*	*-40*	*-20*	*-10*	*0*	*10*	*20*	*40*	*80*
-80	1	1	1	1	1	1	0	0	0
-40	1	1	1	1	1	1	1	1	0
-20	0	1	1	1	1	1	1	1	1
-10	0	1	1	1	1	1	1	1	1
0	0	1	1	1	1	1	1	1	1
10	0	1	1	1	1	1	1	1	1
20	0	1	1	1	1	1	1	1	1
40	0	0	1	1	1	1	1	1	1
80	0	0	0.5	0.5	0.75	1	1	1	1

The vertical coordinate in this table mean the disparity of focused objects; the horizontal coordinate mean disparity of the questioned objects; the value in the table means the fused ability (0:cannot fuse; 1: easy to fuse)

TABLE II The Threshold of Variance (in Terms of Viewing Angle)

DISP	-80	-40	-20	-10	0	10	20	40	80
-80	0.0	-2.1	-3.1	-3.7	-4.2	-4.7	-5.2	-6.3	-8.3
-40	2.1	0.0	-1.0	-1.6	-2.1	-2.6	-3.1	-4.2	-6.2
-20	3.1	1.0	0.0	-0.5	-1.0	-1.6	-2.1	-3.1	-5.2
-10	3.7	1.6	0.5	0.0	-0.5	-1.0	-1.6	-2.6	-4.7
0	4.2	2.1	1.0	0.5	0.0	-0.5	-1.0	-2.1	-4.1
10	4.7	2.6	1.6	1.0	0.5	0.0	-0.5	-1.6	-3.6
20	5.2	3.1	2.1	1.6	1.0	0.5	0.0	-1.0	-3.1
40	6.3	4.2	3.1	2.6	2.1	1.6	1.0	0.0	-2.1
80	8.3	6.2	5.2	4.7	4.1	3.6	3.1	2.1	0.0

The vertical coordinate in this table mean the disparity of focused objects; the horizontal coordinate mean disparity of the questioned objects; the value in the table means the angle difference. Red: cannot fuse; Blue: hard to fuse; Black: easy to fuse

increased from 0. The average view-angle of the first point which is hard to fuse as a plane will be record as the threshold of the visual discomfort.

IV. EXPERIMENT RESULT

In this section, we will show our experiment result to support our proposed models. The three experiments will be discussed separately.

A. Limitation of the View Angle

The experiment result is shown in Fig. 10. The non-convergence region cannot be fused. As our proposed model mentions, when the angle difference is too large, it cannot be fused. The limiting angle difference is about 10.34 degree for the subjects.

B. Variance of theViewing Angle for the Whole Scene

We calculated the corresponding viewing angle variance $|\alpha-\beta|$, where α as the viewing angle of focused object, and β means the viewing angle of the other object. The experimental results are shown in Table I and Table II. We can conclude that the threshold of angle range variance is about 3.6^{o}~4.1^{o} when focusing on the nearest object and it's about 4.7^{o}~5.2^{o} when focusing on the farthest object. However, in a real image, we cannot easily predict where the viewer is interested in. So, we choice 3.6^{o} as the serious upper bound of view angle range variance $|\alpha-\beta|$.

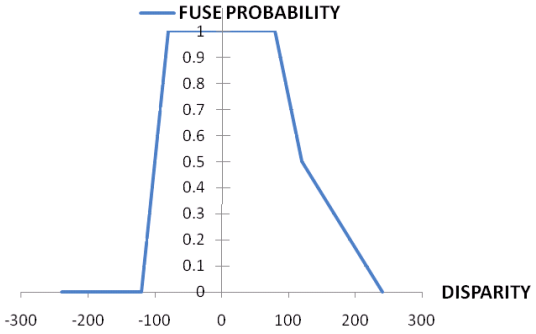

Fig. 10 non-convergence pattern example

C. Inside-object Viewing Angle Differance

The experiment result matches our proposed model. For a given depth, the viewing angle difference can be a good threshold for the fusibility. If the difference is larger than the threshold, it cannot be fused, and vice versa. The further result is shown in Table III. The difference is very small related to the intrinsic angle variance for plate plane patterns.

TABLE III THE RESULT OF EXPERIMENT C: VIEWING ANGLE AND FUSIBILITY

Given Depth (cm)	Angle of Zero Variance (Plate Plane)	Angle of First Hard to Fuse (Threshold)	Difference
100	2.117873	2.119771	0.001898
66.67	3.16481	3.16712	0.00231
57.13	3.687476	3.689938	0.002462
50	4.22429	4.227513	0.003223
44.45	4.726832	4.730391	0.003559
40	5.246208	5.250261	0.004053
33.33	6.280392	6.285182	0.00479
25	8.337736	8.343833	0.006097

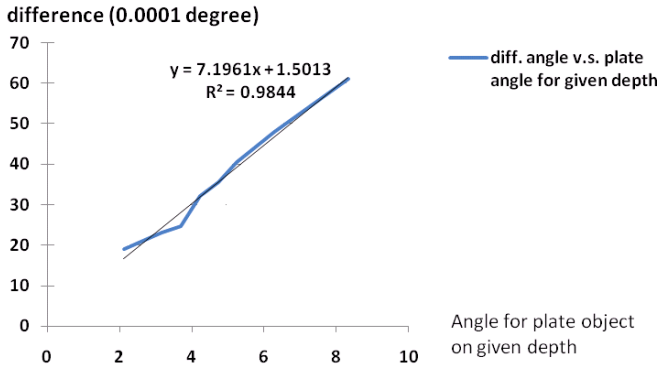

Fig. 11The threshold for angle difference v.s. the angle for a given plate object with fix disparity (depth)

After some mathematical matching, we find the relations about the threshold angle for different given depth in Fig. 11. It's almost linear to the angle for plate object on given depth. The approximated formula for the inside-object viewing angle difference (*IOVAD*) is shown as follows.

$$\mathrm{IOVAD} = 7.2*10^{-5}(Angle(GivenDepth)) + 1.5*10^{-5}. \tag{6}$$

where

$$Angle(GivenDepth) = \tan^{-1}(\frac{EyeDist/2}{GivenDepth})*180^{\circ}/PI. \tag{7}$$

V. CONCLUSION

In this paper, we propose a visual discomfort model for stereoscopic 3D contents. We prove our model by some simplified experiment. According to the proposed model, we can find whether a given stereoscopic 3D content is comfortable for the viewer or not. We believe that this is important for a 3D content provider because to provide a comfortable stereo 3D content for the viewers is important for 3D entertainments.

ACKNOWLEDGMENT

This work is partially supported in part of Himax Technologies, Inc., Taiwan, R.O.C..

REFERENCES

[1] Ianir Ideses, Leonid P. Yaroslavsky, and Barak Fishbain, "Real-time 2D to 3D video conversion," Journal of Real-Time Image Processing, 2007

[2] Donghyun Kim, Dongbo Min, and Kwanghoon Sohn, "Stereoscopic Video Generation Method Using Motion Analysis," IEEE Transactions on Broadcasting 2008

[3] Yong Ju Jung, Aron Baik, Jiwon Kim, and Dusik Park, "A novel 2D-to-3D conversion technique based on relative height depth cue" in SPIE Electronics Imaging, Stereoscopic Displays and Applications XX, 2009

[4] H. Murata et al.: "A Real-Time 2-D to 3-D Image Conversion Technique Using Computed Image Depth", SID Digest of Technical Papers, 32.2, pp919-922 (1998)

[5] T. Iinuma, H. Murata, S. Yamashita, K. Oyamada, "Natural Stereo Depth Creation Methodology for a Real-time 2D-to-3D Image Conversion," SID Symposium Digest of Technical Papers, 2000..

[6] Royal Philips Electronics, "Whitepaper WOWvx BlueBox" [Online]. Available: http://www.business-sites.philips.com/global/en/gmm/images/3d/3dcontentceationproducts/downloads/BlueBox_white_paper.pdf.

[7] C. Wu, G. Er, X. Xie, T. Li, X. Cao, and Q. Dai, "A novel method for semi-automatic 2D to 3D video conversion," 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2008, pp. 65–68, May 28–30, 2008.

[8] H. M. Wang, C. H. Huang, and J. F. Yang, "Block-based depth maps interpolation for efficient multiview content generation," IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 12, pp. 1847–1858, Dec. 2011.

[9] Y. Nojiri, H. Yamanoue, S. Ide, S. Yano, and F. Okano, "Parallax distribution and visual comfort on stereoscopic HDTV,"In Proceedings of IBC, pp.373-380, 2006

[10] Hosik Sohn, Yong Ju Jung, Seong-il Lee, Hyun Wook Park, Yong Man Ro, "Attention model-based visual comfort assessment for stereoscopic depth perception," *Digital Signal Processing (DSP)*, pp.1-6, 6-8 July 2011

[11] Christian Richardt and Lech Ŝwirski and Ian Davies and Neil A. Dodgson, "Predicting Stereoscopic Viewing Comfort Using a Coherence-Based Computational Model," *Proceedings of Computational Aesthetics*, August 2011