Perceived Depth Modeling Based on Color Information

Jeong-Yeop Kim, Sang-Hyun Kim, and Gil-Ja So, *Member, IACSIT*

Abstract—**Conventional perceived depth models represent the perception of depth by human being considering algebraic structure of observation environment. In this paper, we want to propose a new perceived depth model using the fact that the perceived depth can be varied with the change of color components of observation contents while the observation environment fixed.**

Index Terms—**Color component, perceived depth, human visual system.**

I. INTRODUCTION

3D image is now generalized everywhere especially on 3D movie theater, TV, video and so on. However, there is no good guideline or some reference data for making these images or editing them. This is because there are so many variables in capturing images including illuminant, capture angle for camera, registration problem for two left and right camera and hardware characteristic for capturing system. And there is another problem when the observer watches 3D images or video on some display.

The characteristic of display device and observation condition such as distance from observer to device, size of display, dot pitch of display, inter-pupil distance of observer, disparity of 3D image and so on. Actually, all these parameters determine the quality of 3D image on display when the observer watches the contents. There are many ways to compensate these parameters to give more realistic depth and image qualities to observers. Usually, professional 3D contents are optimized for particular observation condition, the characteristics of capturing device and so on.

However, UCC (User created contents) usually have no device calibration, correction for observation condition and parameters for observer. Because of having no specific technical information about the environment, general users can not calibrate their captured images and there's no way for them to adjust their images at all. In this case, almost all parameters regarding optimal observation environment including visual discomfort zone[1] and [2] are ignored and therefore gives some visual fatigue when observer watch 3D contents[3] and [4]. Visual discomfort zone means that if the observer is in discomfort zone when he or she observes 3-D images on some display, he or she feels discomfort. In other case, when the observer is in comfort zone, he or she feels no

discomfort for the 3-D images. This phenomenon comes from the geometrical parameters such as the distance between the observer and the display, the inter-pupil distance of the observer, disparity of the image pixel position and dot pitch of display device and so on. This feeling of discomfort causes fatigue when the observer watches 3-D images. The fatigue does not caused from only these parameters mentioned above, however these parameters has potential influence on it. Because there are so many other reasons for the observer feels fatigue. For example, the physical conditions, emotional changes, and so on. In the researches related the perception of depth assumes the fatigue comes from mainly from the geometrical parameters mentioned above and usually ignores those other reasons. In this paper, we also follow this assumption as well.

One of the methods includes depth tuning technique by defining safety zone. ChangYuan [5] *et al*. calculated this kind of safety zone in some particular observation environment with fixed size of display device. They controlled the depth information by shifting and scaling based on human visual characteristics.

In this paper, we are interested in representation of perceived depth information rather than tuning or modifying the disparity data directly and a new perceived depth model is proposed by using this phenomenon. In the near future, we can adjust perceived depth information by only adjust the color values and luminance rather than disparity itself.

II. CONVENTIONAL METHOD

A. Conventional Perceived Depth Model

Perceived depth model has been considered by many researchers and most of them did not take the effect of human visual factor and just thought about the algebraic structure of observation environment.

According to the research of Chang Yuan [5] *et al*, if the stereo images are captured in parallel mode camera, the perceived depth is proportional to disparities and can be represented in (1) where, D is perceived depth, V is observation distance (distance from observer to display device), I is IPD (inter-pupil distance) of observer, S_D is horizon pixel pitch of display device and d is disparity information. The visual comfort zone is computed as in Fig. 1.

$$D = \frac{V}{\frac{I}{S_D d} - 1} \tag{1}$$

In Fig. 1, the main parameters to decide visual comfort are observer-to-screen distance and disparity at screen. These two parameters have great influence on fatigue.

They shifted the images laterally to adjust disparity image and correct the data outside visual comfort zone. Additionally, if the shift values are not compatible for the scene, for

example, very big to be shifted, these data can be scaled. By this approach, they can calibrate the 3-D images and make it comfortable when the observer watches the images. But there exist some artifact for disparity information because this approach adjusts the disparity values directly and there can be some discontinuity by this adjustment. In this paper, we are interested in some method how can we adjust the degree of depth not by direct adjustment of disparity.
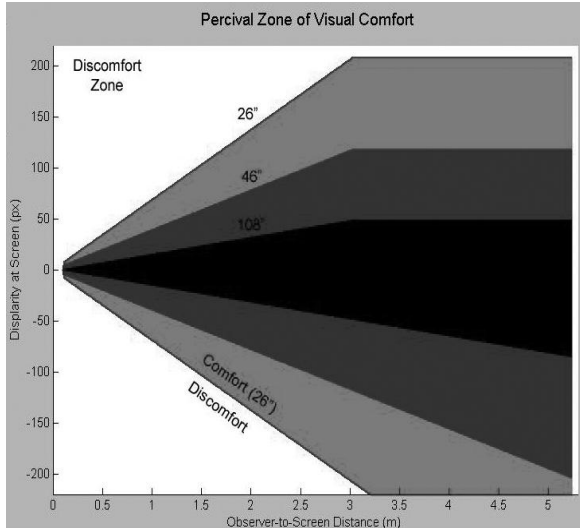

Fig. 1.Visual comfort zone for various kind of display.

Perceived depth model represents the actual depth perceived for some 3D images by human being in some observation condition. If we fix the observation condition such as observation distance, display device and so on, the perceived depth model gives constant depth information regardless of the displayed contents [5]. It is not considered the property of contents displayed on display device and the perceived depth is represented by just the physical, algebraic parameters and proportional rate of observation environment.

However, according to the human visual experiment of depth perception [6], the actual perceived depth can be varied with the change of displayed contents such as wavelength or luminance values.

B. Human Visual Experiment on Depth Perception

However, it is known that hue and luminance components give some influence on depth perception of human-being from the physiological experimental reports about human vision and depth perception.

Cary Lu [6] *et al* set experiment environment like Fig. 2 and generated stereopsis images like Fig. 3. Stereopsis image by many type of combination was generated and used in experiment on if human being can recognize the depth information for the images or not. The contents of combination can be represented as wave length (λ) of patch luminance (*L*) information. These two factors can be manipulated by experimenter and therefore, gives many patterns for this experiment.

In Fig. 2, they made an analog experiment environment by using optical instruments such as telescopes, optical filters, physical patterns, photo multipliers, and light sources. By deploying these devices as shown in Fig. 2, they can make some optical stereogram as shown in Fig. 3. There are some reflection, refraction and optical filtering to make the optical
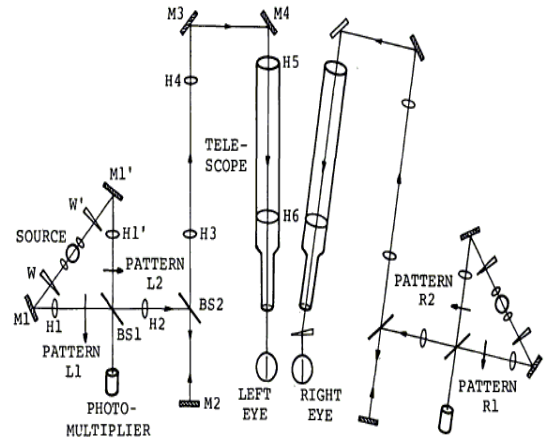
images in Fig. 3.

Fig. 2. System for optical construction of stereopsis images.
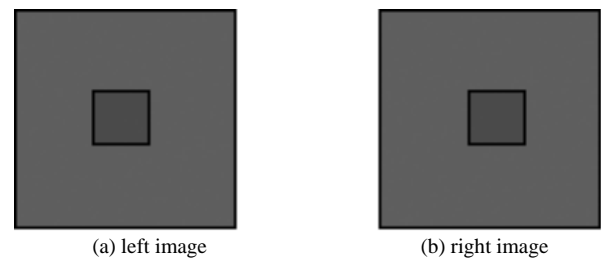

| (a) left image | (b) right image |

Fig. 3. Stereopsis image by optical construction.

By the above experiment environment, they generated some patterns by changing the wavelength and luminance of red and green patches. The result is shown in Table I and Table II. The depth of green object on red background is well perceived at low luminance (average contrast 10.5). The other case, red object should have high luminance (average contrast 1.5) to be perceived well. This fact of hysteresis means hue and luminance gives some influence on depth perception of human being.

TABLE I: RESULT OF CONTRAST: $\lambda_1 = 634nm$, $\lambda_2 = 512nm$

Patch1(red)	Patch2(green)1	Contrast ratio(Patch 1/Patch 2)2
256.47	26.40	9.78
167.25	13.00	12.90
104.25	8.95	11.68
54.40	4.62	11.75
34.21	3.86	8.80
22.12	2.42	9.47
13.08	1.42	9.17

In this paper, we want to propose the perceived depth model based on this fact considering wavelength and luminance of displayed contents.

TABLE II: RESULT OF CONTRAST: $\lambda_1 = 512nm$, $\lambda_2 = 634nm$

Patch1(red)	Patch2(green)1	Contrast ratio(Patch 1/Patch 2)2
199.44	151.16	1.32
97.36	45.38	2.17
69.00	45.08	1.53
45.03	31.33	1.45
30.20	22.15	1.37
19.30	15.03	1.27
11.91	7.89	1.48

III. EXPERIMENTAL ENVIRONMENT

The test synthetic images for this new model can be generated by 3dhippie stereo-cam v1.6.4 [7] of Fig. 4 and be shown in Fig. 5.
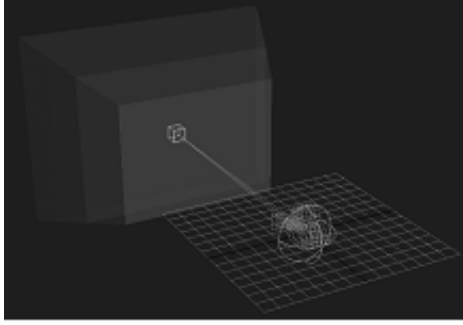

Fig. 4. Stereo-Cam plugin for 3Ds-Max.

(a) left image   (b) right image

Fig. 5. Test stereopsis image by plugin.

IV. PROPOSED MODELING

The information such as size of red and green patches, distance between camera and patches, wavelength and luminance can be obtained from the test environment by plugin. This image matches to Table I of Cary Lu *et al*. L_1 means luminance of patch1 and λ_1 is wavelength. The same for patch2 like L_2 and λ_2. According to Table I, the depth can be recognized at luminance contrast (L_1/L_2) is larger than 9.78 and the other case, the depth cannot be perceived. In this case, the perceived depth information of patch2 can be represented in some value by (1). The actual perceived depth for this case is 0 and human being cannot be perceived the difference of depth between patch1 and patch 2. Therefore, the wavelength and luminance of patch2 gives some influence on depth perception. Equation (1) can be modified as (2) based on this fact.

$$D_c = \frac{V}{\frac{I}{S_D d}-1} \times f(\lambda_1, L_1, \lambda_2, L_2) \qquad (2)$$

The constraint can be found in Table I and Table II. From Table I, D_c should be 0 when the luminance contrast is 9.78 for $\lambda_1 = 634nm$, $\lambda_2 = 512nm$. From Table II, D_c should be 0 when the luminance contrast is 1.32 for $\lambda_1 = 512nm$, $\lambda_2 = 634nm$. Therefore, the new model should satisfy these two constraints simultaneously.

The constraints are shown in (3) as follows:

$$D_c = \begin{cases} 0 \ if \ \lambda_1 = 634, \lambda_2 = 512, \ \frac{L_1}{L_2} = 9.78 \\ 0 \ if \ \lambda_1 = 512, \lambda_2 = 634, \ \frac{L_1}{L_2} = 1.32 \end{cases} \qquad (3)$$

The function in (2) can be represented as (4) to accommodate these constraints. If we substitute the values of λ_1, λ_2, L_1, L_2 into (2) and left side would be zero for these two cases.

$$f(\lambda_1, L_1, \lambda_2, L_2) = a \times \left(\frac{L_1}{c_1\lambda_1 + c_2\lambda_2} - L_2 \right) \qquad (4)$$

The value of constraints applied to (4) and (5) is obtained where a is a proportional constant to adjust the difference between the perceived and actual depth.

To solve this constraint, we made some graphical representation for (4) as shown in Fig. 6.
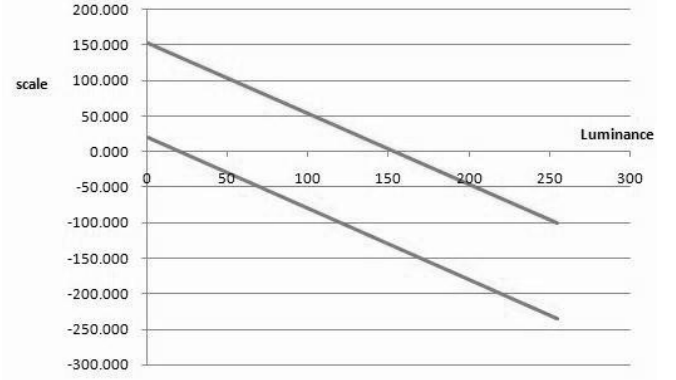

Fig. 6. Scale values for functional.

The lower line comes from the first raw of Table I, and the upper line comes from the first raw of Table II. If (4) satisfies the constraint in (3), the scale values of vertical axis gives the exact value of functional (4). After some mathematical operations, the constants c_1 and c_2 can be calculated as in (5).

$$f(\lambda_1, L_1, \lambda_2, L_2) = a \times \left(\frac{L_1}{0.0396\lambda_1 - 0.0299\lambda_2} - L_2 \right) \qquad (5)$$

Now, the proposed model can be represented in (6).

$$D_c = \frac{V}{\frac{I}{S_D d}-1} \times a \times \left(\frac{L_1}{0.0396\lambda_1 - 0.0299\lambda_2} - L_2 \right) \qquad (6)$$

The actual and perceived depth information can be calculated and shown in Fig. 7. The horizontal axis of Fig. 7 represents the horizontal coordinates of the experimental image and gives a kind of cross section data of center position.

There is another experimental result as input image shown in Fig. 8. The image consists of red ball on the green background and this image is different from images from Fig. 5. Because the image of Fig. 5 has a planar shape and the luminance of red part is the same, however the luminance of red ball of Fig. 8 is different according to the position of pixels. In Fig. 9(a), RGB values of the cross section in center part of Fig. 8(a) are presented.

The synthetic image generated has sphere shape and therefore the proposed model does not be applied directly because the depth information is different with the change of pixel position. And the pixel value i.e. luminance value is also varies in many cases with the change of incoming angle of illuminant source and the geometry of objects position.
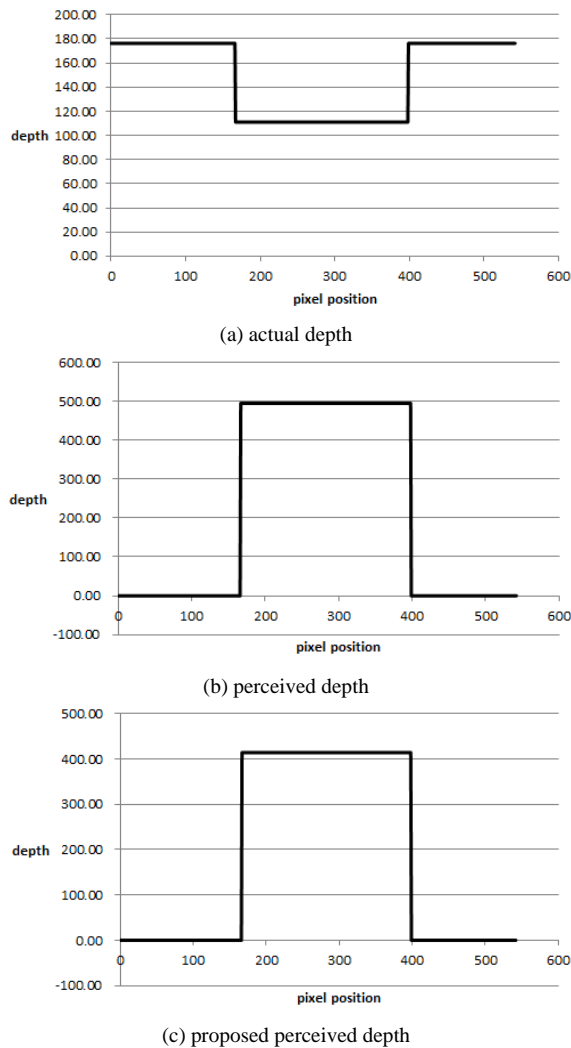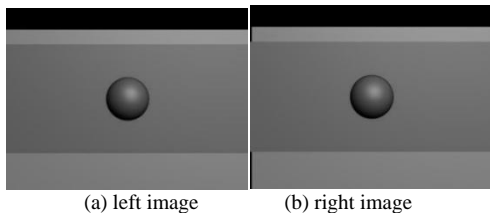
(a) actual depth

(b) perceived depth

(c) proposed perceived depth

Fig. 7. Profile information of image.

(a) left image   (b) right image

Fig. 8. Stereopsis image by optical construction.
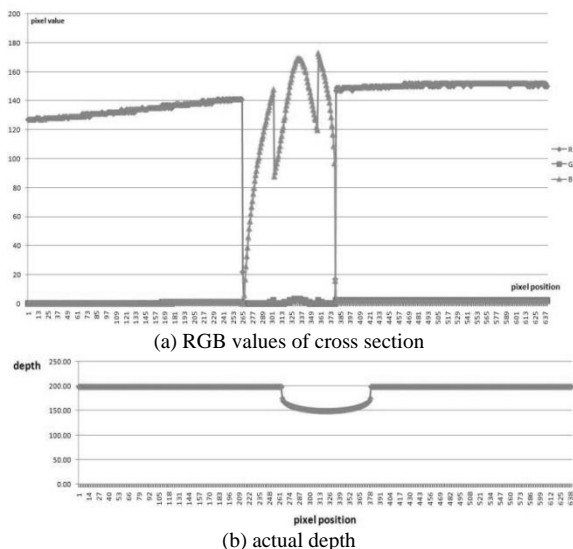

(a) RGB values of cross section

(b) actual depth

Fig. 9. Profile information of synthetic image.

V. CONCLUSION

In this paper, the method for finding relationship between the change of saturation values and perceived depth information was proposed. By the proposed method, we found if we decrease saturation values for some object in the scene, it looks like more farther than its real location. This fact can be utilized in reducing the viewing fatigue for 3D stereoscopic images because reduce in saturation values to some extent causes the change of perceived depth information without any distortion in depth. Therefore, the viewing fatigue can be controlled in some extent without distortion of depth information. The proposed perceived depth model can be modified to satisfy the conditions of real environment.

REFERENCES

[1] L. Marc, F. Marten, H. Ingrid, and I. Wijn, "Visual discomfort and visual fatigue of stereoscopic displays: A review," *Journal of Imaging Science and Technology*, vol. 53, no. 3, pp. 1-14, May 2009.

[2] S. Lee, Y. Jung, H. Sohn, Y. Ro, and H. Park, "Visual discomfort induced by fast salient object motion in stereoscopic video," in *Proc. SPIE 7863, Stereoscopic Displays and Applications XXII 786305*, San Francisco, California, pp. 1-2, 2011.

[3] I. P. Howard and B. J. Rogers, "Binocular vision and stereopsis," *Oxford University Press*, ed. 1, 1995.

[4] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-Accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of Vision*, vol. 8, no. 33, pp. 1-30, 2008.

[5] C. Yuan, H. Pan, and S. Daly, "Stereoscopic 3D content depth tuning guided by human visual models," in *Proc. SID Symposium Digest of Technical Papers*, vol. 42, issue 1, pp. 916-919, June 2011.

[6] C. Lu and D. H. Fender, "The interaction of color and luminance in stereoscopic vision," *Investigative Ophthalmology*, pp. 484-490, June 1972.

[7] 3dhippie-stereocam. [Online]. Available: http://www.scriptspot.com/3ds-max/scripts/3dhippie-stereocam

Jeong Yeop Kim received the Diploma in electronics engineering from Kyung-Pook National University of Daegu, South Korea in 1990, and MS and PhD degrees, both in image processing of electronics engineering, from the same university, in 1992 and 2000, respectively. He was with the Samsung Advanced Institute of Technology in Korea as a senior research staff from 1994 to 2000. From March 2001, he is with the Department of Game Contents Science of the University of Young-San at Busan, as an associate professor. His doctoral dissertation work focused on detecting and adjusting the white point from single color image. The topics he has worked on include color printing, error diffusion, color correction and 3D stereoscopic images. His current research is on correction of color data in 3D stereoscopic images. He is a member of the KMMS (Korea Multimedia Society) and KIPS (Korea Information Processing Society).

Sang Hyun Kim received the Diploma in electronics engineering from Kyung-Pook National University of Daegu, South Korea in 1991, and MS and PhD degrees, both in image processing of electronics engineering, from the same university, in 1993 and 1998, respectively. He was with the Samsung Medison Co. Ltd, in Korea as a senior researcher from 1998 to 2001. From September 2001, he is with the Department of Computer Engineering of the University of Young-San at Yangsan, Kyungnam as an associate professor. His doctoral dissertation work focused on the development of fractal image coding method and transmission for image processing, image communication and computer vision problems. The topics he has worked on include virtual reality, ultrasound image processing, medical image processing, image feature extraction and classification, and 3D stereoscopic images. His current research is on region-based intensity correction in 3D stereoscopic images. He is a member of the KMMS (Korea Multimedia Society) and KIPS (Korea Information Processing Society).

Gil Ja So received the Diploma in computer science from the Dong-eui University of Pusan, South Korea in 1994, and MS and PhD degrees, both in natural language processing of computer engineering, from the Pusan National university, South Korea in 1997 and 2012, respectively. From September 2001, she is with the Department of cyber police and science of the University of Young-San at Yangsan, Kyungnam as an assistant professor. Her doctoral dissertation work focused on the Korean Grammar Checker for parsing Korean documents and correcting context-sensitive words in documents. The topics she has worked on include virtual reality, game development and artificial intelligence of non-players of game, and 3D stereoscopic images. Her current research is on region-based intensity correction in 3D stereoscopic images. She is a member of the KMMS (Korea Multimedia Society) and KIPS (Korea Information Processing Society).