3-D Perception Enhancement in Autostereoscopic TV by Depth Cue for 3-D Model Interaction

Yi-Ting Shen*, Guan-Lin Liu†, Sih-Sian Wu‡, and Liang-Gee Chen§
Department of Electrical Engineering
National Taiwan University, Taipei, Taiwan
DSP/IC Design Lab
Email: b01901100@ntu.edu.tw*, b00901064@ntu.edu.tw†, benwu@video.ee.ntu.edu.tw‡, lgchen@video.ee.ntu.edu.tw§

***Abstract*—Autostereoscopic TV provides users 3-D experience without wearing glasses. To overcome its poor resolution and to improve the 3-D perception, in this paper, based on Human Visual System (HVS) theory, some depth cues were added into the proposed system. It is shown in the result that the viewers can perceive the 3-D models twice as much as the conventional one. A 3-D model interaction system was built with the proposed framework.**

Keywords: autostereoscopic TV, depth cue, 3-D interaction

I. Introduction

Autostereoscopic three-dimensional(3-D) television (3DTV) is a developing trend in the display industry, from which consumers can experience the same 3-D perception effect without wearing glasses or other devices. With this advantage, autostereoscopic TV can be widely used in business advertisement, since it not only displays the product's 3-D model, but also allows users to touch the 3-D images. However, in spite of the attractive functions, autostereoscopic TV is still far away from popular due to three major problems.

A. Common Problems

Firstly, in order to give users multiple angles of view to form 3-D images, automultiscopic screens present different images depending on the viewing direction (Fig. 1). The screen of the TV has been divided into several vertical cuts, sending distinct and directional image slots. In other words, the spatial resolution has been dragged down due to this technique. Secondly, since the limited angular resolution of such display, viewers might find some vertical distortion caused by view transition at specific angles [1]. This might make people feel uncomfortable to stand in these regions. Thirdly, the resulting 3-D effect is still not good enough. In our experiment, viewers could at most feel the image models that we built at the position about 5 cm in front of the screen. This narrow visual comfort region of autostereoscopic TV may become the most significant restriction to its application [2].

B. Our Work

In this paper, an alternative solution for autostereoscopic TV without modifying the device is proposed. Due to the restriction of low spatial resolution, a 3-D model display system is constructed. This system is not used for ordinary TV applications, which requires higher resolution for small words and images. In addition, to interact with 3-D image models, users should be able to touch the models when they stand in the clear view region, which is at least 1 m in front of the TV. To achieve this goal, we added some depth cues to enhance the 3-D perception of autostereoscopic TV. Comparison and numerical analysis on each of the cue was made.

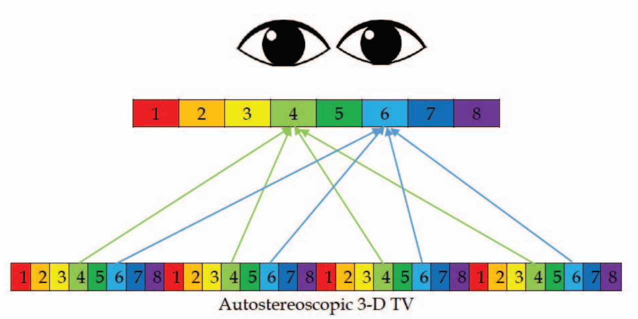

Fig. 1. A demonstration of display principle for autostereoscopic 3-D TV. The viewer's eyes perceive different view depending on their viewing direction. In this case, the viewer's right eye perceive the view 4, and the other eye perceive the view 6.

The paper is organized as described in the following. In section two, related work is listed and briefly introduced. Our experimental methods and the proposed system are presented in section three. Section four contains the experimental results and the analysis on each depth cue. Section five highlights the applications of this work and concludes the outcome of this paper.

II. Related Work

A. Other Solutions

Many other solutions have been proposed to improve autostereoscopic TV displays [1, 3]. However, most of these solutions need to modify the autostereoscopic 3-D TV hardware, or add extra image processing algorithm, which might increase the cost and thus make it difficult to become a commercial product. In view of this, a new technique without changing any original display devices is proposed in this paper.

B. Depth Cues

Based on the Human Visual System (HSV) theory, human perception of depth is composed of the optical perception through the eyes and the brain operation [4]. The brain operation involves complex reasoning which is related to the experiences, such as texture, occlusion, related size, cast shadow and motion parallax caused by the movement of the viewer's position. According to the HSV theory, we assumed that some depth cues may be useful for enhancing the 3-D perception effect (Fig. 2).

Much research has been conducted on autostereoscopic mobile displays and the results has proved that adding depth cues is an effective technique to enhance the 3-D perception [9, 10].

However, for autostereoscopic 3-D TV displays, as the distance between the monitor and the users becomes larger, the viewing direction becomes wider as well. In addition, the displays should be available for multiple users. All of these problems may not happen in autostereoscopic mobile display application so as to be taken into consideration additionally.

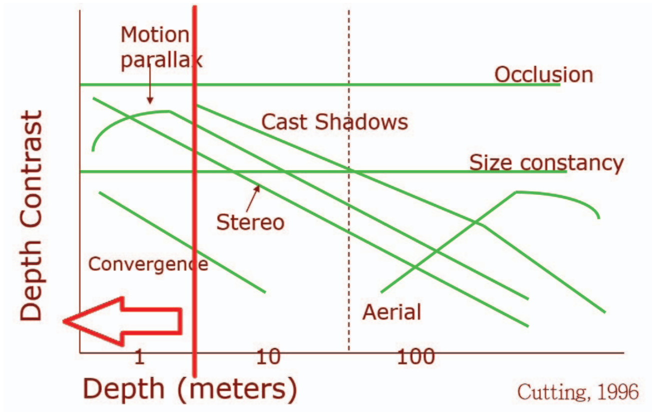

Fig. 2. The relationship between sensitivity toward each depth cue and distance. Humans have different sensitivity toward each depth cue, and the sensitivity is different distance by distance [5]. In the proposed interaction system, the regions below 1 m are our main research area [6].

C. Motion Parallax

As the viewers move, objects that are closer to their move farther across our field of view than objects that are in the distance. To have a better 3-D perception effect, viewers should receive different angles of view when they are moving. In addition, geometric distortions will occur on multi-view displays with only horizontal parallax [7]. It is shown that the ordinary motion parallax effect provided by automultiscopic screens is not complete. The traditional motion parallax technique, frequently used in 2-D to 3-D conversion system [8], is applied in our proposed system to provide viewers with a better freedom of position.

III. Proposed System and Experimental methods

A. Proposed System

The proposed system is shown in Fig. 3.

B. Experimental Methods

Based on General Unified Theory [5]:

$$perceived\ depth = weighted\ sum\ of\ all\ depth\ cues \quad (1)$$

In our experiments (based on Fig. 2), six different kinds of depth cues were adjusted reasonably. Three human subjects have been invited to help us conduct three different kinds of experiments.

In the first experiment, six depth cues were individually added to the view and the subjects were asked to touch the 3-D image model. We measured the distance between the TV and the position that the subjects felt the model being. The inequality of each depth cue's effect was proved and their strength was defined by a numerical method.

In the second experiment, two scenes were given to the subjects at the same time (Fig. 6-10). Both of them were enhanced by the same kind of depth cue, but in different strength. Our goal was to find the effect of adjusting each parameter on the depth perception. Firstly, the subjects were asked which one they

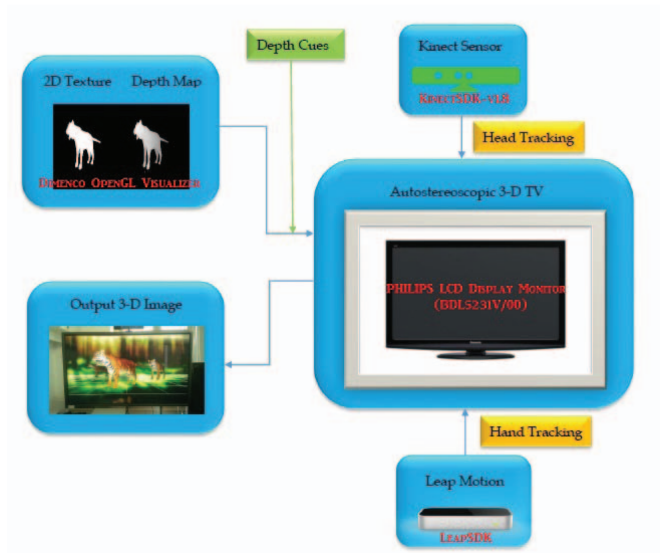

Fig. 3. Purposed System. The LCD Display Monitor is used to combine a 2-D texture with a depth map to form a 3-D image. The Kinect Sensor is used to capture viewer's head position, and the Leap Motion Sensor is used to collect the information of viewer's hands. The open source language (openGL) is used to create the scene, which is shown by Dimenco OpenGL Visualizer.

thought was better than the other simply based on their feelings. Secondly, similar to the first experiment, they were asked to touch the models. From the acquired data, it was found that even the viewers said one scene is better than the other or equal to the other, the result of the measurements are sometimes different. This is mainly because human's brain cannot be precise enough for the comparison in such a small distance. The two sets of data should be combined and analyzed together to figure out the final result.

In the last experiment, based on the experiments indicated in the above, all the depth cues were combined, and the best and most comfortable values for each parameter were found. According to our experiment, a better 3-D effect can be obtained if all the strongest depth cues are combined together. However, if the best 3-D effect is desired, the whole view should be made as real as possible, which is easier for our brain's depth operation. The six parameters (depth cues) are listed and explained below.

1) Brightness of texture: Texture presents the correlation between the model and the environmental light. The image processing program (Adobe Photoshop CS5.1) was used to adjust the texture attached to the model. The brightness was adjusted between -40 to 40 unit level in the Adobe Photoshop CS5.1.

2) Relationship between models: The relationship between models is included in the "size constancy" and "occlusion" in Fig. 2. Another model was put into the scene with different distances from the original model. The distance between the models was adjusted between 2 to 4 unit length in the openGL.

3) The density of linear perspective cue : This is included in the "convergence" in Fig. 2. The same distance between two lines looks shorter when they are getting farther. Different strength of perspective cue by adjusting the distance between each line was given. The distance was adjusted between 0.1 to 1 unit length in the openGL.

4) The size of cast shadow: This is included in the "cast shadow" in the Fig. 2. With the same position of light and model, the size of shadow gives a hint of the distance between the

978-1-4673-8364-6/16/$31.00 ©2016 IEEE

model and the background. The size of the shadow was adjusted between 0.96 to 1.08 times of the model.

5) The strength of camera lens focus effect: This is included in the "convergence" in Fig. 2. When viewers focus on the near-scene, the far-scene should be out of focus (blurring). Different Gussian blur strength was given to the original background picture in the Adobe Photoshop CS5.1. The Gussian radius was adjusted between 0.1 to 8.0 pixels.

6) With and without the extra motion parallax effect: This is included in the "motion parallax" in Fig. 2. After constructing the whole 3-D scene in the openGL, what the viewers have seen is controlled by the position which the camera focuses on. To apply the motion parallax effect, the viewport is changed according to the head position data fetched by a Kinect sensor based on the formula below :

$$\begin{aligned} displacement\ of\ the\ camera &= \\ displacement\ of\ the\ viewer's\ head/10 \end{aligned} \quad (2)$$

IV. Experimental Results

A. Experiment I

Depth Cue / Subject	A (cm)	B (cm)	C (cm)
Default	5	6.5	7
Texture	7	7	8
Linear Perspective Cue	7	7	9
Relationship between models	9.5	7	14
Depth of Focus	10.2	7.1	9
Cast Shadow	7	7	9

TABLE I
Experiment I

According to Table I, it seems that "Relationship between models" and "Depth of Focus" are stronger than the other parameters. In addition, according to the Table II, viewers can always have better 3-D perception if the textured model is used with other depth cues. As a result, model with texture was used in the following experiments.

Depth Cue / Subject A	Texture-free (cm)	Textured (cm)
Linear Perspective Cue	7	7.1
Relationship between model	9.5	10
Depth of Focus	10.2	11.2
Cast Shadow	7	8.5

TABLE II
Experiment I : Comparison between textured model with texture-free model

B. Experiment II

In Table III to Table. VIII, "similar" indicates that the subject felt it as close as "Middle one" does, "in" indicates that the subject felt it farther than "Middle one" does, "out" indicates that the subject felt it closer than "Middle one" does, and "strange" means that the model made the subject feel uncomfortable.

1) Texture: From the Table. III, it is shown that brighter texture can make the viewers feel the model being closer to them.

Subject	Dark (cm)	Middle (cm)	Bright (cm)
A	7 (in)	11	11 (similar)
B	5 (in)	6	5.2 (out)
C	3 (in)	7	11 (out)

TABLE III
Experiment II : Model with bright or dark texture.

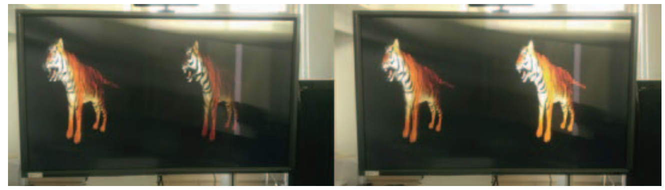

Fig. 4. With bright and dark texture.

2) Relationship between models: According to Table IV, if the models are too close, the viewers can't sense the distance between them and will have a bad 3-D perception.

Subject	Near (cm)	Middle (cm)	Far (cm)
A	11.5 (similar)	12	12 (similar)
B	5.2 (in)	6.8	3.8 (out)
C	4.5 (strange)	4.5	3 (out)

TABLE IV
Experiment II : With another close or far model.

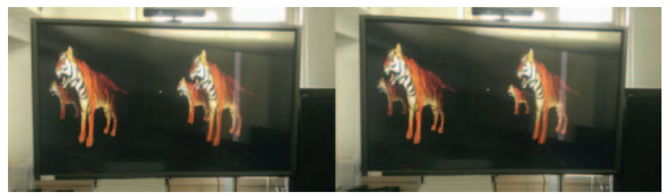

Fig. 5. With another close or far model.

3) Depth of Focus: From the Table V, the level of blurring should be reasonable, or the viewers might feel uncomfortable.

Subject	Blur (cm)	Middle (cm)	Clear (cm)
A	12 (in)	12	11.5 (strange)
B	4.2 (in)	7	4.2 (in)
C	8 (strange)	3.5	4.5 (strange)

TABLE V
Experiment II : With blur or clear background.

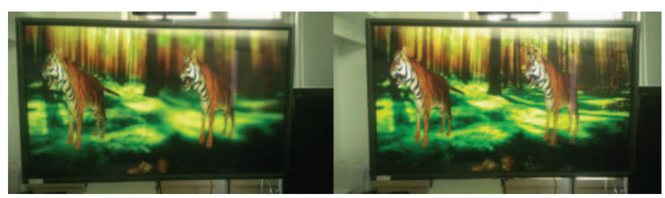

Fig. 6. Create different level fake camera lens effect.

4) Cast Shadow: According to Table VI, the size of shadow should be suitable to the model, so that viewers may not feel it strange.

Subject	Small (cm)	Middle (cm)	Large (cm)
A	11(in)	11	11(in)
B	2.5(in)	6.4	4.1(similar)
C	8(strange)	4	3(strange)

TABLE VI
Experiment II : With different size of shadow attached to the model.

5) Linear Perspective Cue: Viewer's sensitivity toward linear perspective cue can be extremely different (Table VII). However, it will make the whole scene more realistic, especially when combining with "motion parallax" effect.

978-1-4673-8364-6/16/$31.00 ©2016 IEEE

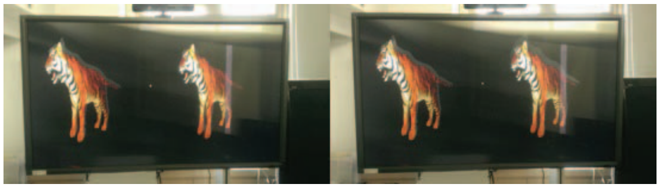

Fig. 7. Attach different-size shadow to the model.

Subject	Dense (cm)	Middle (cm)	Thin (cm)
A	13.5 (similar)	13.5	13.5 (in)
B	4.4 (similar)	7	4 (similar)
C	3 (out)	2.5	1.5 (out)

TABLE VII
EXPERIMENT II : WITH DIFFERENT DISTANCE BETWEEN LINES.

Fig. 8. Create the ground with linear perspective cue.

6) Motion Parallax: Apart from the original motion parallax effect provided by the autostereoscopic TV, the new motion parallax effect can allow up and down motion, from which the subjects felt the scene more realistic.

C. Experiment III

According to experiment I and II, some design rules for the 3-D perception enhancement are found : (1) The brighter model looks closer. (2) The 3-D perception of near-scene is enhanced by linear perspective ground and adding another model with different depth and texture. (3) The 3-D perception of far-scene is enhanced by adding a blurred background. (4) Applying modified motion parallax effect can get a better 3-D perception.

Combining all depth cues based on these rules (Fig. 9), the viewers depth perception is enhanced up to 4 times as much as before. The result is shown in Table VIII.

Subject	A(cm)	B(cm)	C(cm)
Before	5	4.8	2.5
After	14	7.1	9.5

TABLE VIII
EXPERIMENT III : LINEAR COMBINATION OF ALL DEPTH CUES.

Fig. 9. Linear combination of all depth cues.

V. APPLICATIONS AND CONCLUSION

A. Applications

With the aid of leap motion sensor, the position of user's hands and fingers can be detected, so that viewers can control the models by specific gesture (Fig. 10). For practical applications, since the autostereoscopic TV is not popular, it will promote product advertisement if it is placed on the street. In education, we may apply the technique to create a virtual 3-D museum. People can touch and learn the knowledge of some masterpieces or endangered species which only exist at some particular places in the world.

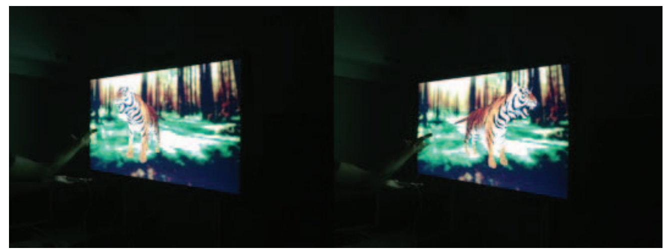

Fig. 10. control the model by leap motion.

B. Conclusion

In this work, an effective way to strengthen the 3-D perception of autostereoscopic TV by simply adding some depth cues in the view is proposed. In this way, a 3-D model interaction display monitor is made. However, more complicate depth cues were not considered, thus the 3-D effect that we have achieved is still far away from perfection. In the future, more depth cues will be taken into consideration and more accurate analysis on the relationship between the depth cues will be performed. It is believed that enjoying the best 3-D experience without wearing any device is around the corner.

REFERENCES

[1] Song-Pei Du, Piotr Didyk, Frédo Durand, Shi-Min Hu, and Wojciech Matusik. "Improving Visual Quality of View Transitions in Automultiscopic Displays." ACM Transactions on Graphics (Proceedings SIGGRAPH Asia 2014, Shenzhen, China), 2014.

[2] Lambooij, Marc TM, Wijnand A. IJsselsteijn, and Ingrid Heynderickx. "Visual discomfort in stereoscopic displays: a review." Electronic Imaging 2007. International Society for Optics and Photonics, 2007.

[3] Woodgate, Graham J., et al. "Autostereoscopic 3D display systems with observer tracking." Signal Processing: Image Communication 14.1 (1998): 131-145.

[4] Sung-Kyu Kim, Dong-Wook Kim, Yong Moo Kwon, and Jung-Young Son. "Evaluation of the monocular depth cue in 3D displays," Opt. Express 16, 21415-21422 (2008).

[5] De Silva, Varuna, et al. "Sensitivity analysis of the human visual system for depth cues in stereoscopic 3-D displays." Multimedia, IEEE Transactions on 13.3 (2011): 498-506.

[6] Matt Williams. "Depth perception and visualization".

[7] Amir Said and Bruce Culbertson. "Analysis and management of geometric distortions on multi-view displays with only horizontal parallax." HP Laboratories. HPL-2012-2.

[8] Matsumoto, Yukinori, et al. "Conversion system of monocular image sequence to stereo using motion parallax." Electronic Imaging'97. International Society for Optics and Photonics, 1997.

[9] Mikkola, Maija, Atanas Boev, and Atanas Gotchev. "Relative importance of depth cues on portable autostereoscopic display." Proceedings of the 3rd workshop on Mobile video delivery. ACM, 2010.

[10] Huhtala, Jussi, et al. "Evaluating depth illusion as method of adding emphasis in autostereoscopic mobile displays." Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services. ACM, 2011.

978-1-4673-8364-6/16/$31.00 ©2016 IEEE