A Technique for Precise Depth Representation in Stereoscopic Display

SHUNSUKE YOSHIDA*1, SHINYA MIYAZAKI*2, TOSHIHITO HOSHINO*3, TORU OZEKI*3, JUNICHI HASEGAWA*2, TAKAMI YASUDA*4 and SHIGEKI YOKOI*1

**1 Graduate School of Human Informatics, Nagoya University*

Furo-cho, Chikusa-ku, Nagoya-city, Aichi, 464-8601, Japan

shun@info.human.nagoya-u.ac.jp

yokoi@info.human.nagoya-u.ac.jp

**2 School of Computer and Cognitive Sciences, Chukyo University*

Tokodachi 101, Kaizu-cho, Toyota-city, Aichi, 470-0393, Japan

miyazaki@sccs.chukyo-u.ac.jp

hasegawa@sccs.chukyo-u.ac.jp

**3 Toyota Motor Corporation*

Toyota-cho 1,Toyota-city, Aichi, 471-8572, Japan

hoshino@design.en.toyota.co.jp

ozeki@design.en.toyota.co.jp

**4 School of Informatics and Sciences, Nagoya University*

Furo-cho, Chikusa-ku, Nagoya-city, Aichi, 464-8601, Japan

yasuda@info.human.nagoya-u.ac.jp

Keywords: stereoscopic display, depth perception, virtual reality, registration

Abstract

In observing a virtual 3D object displayed stereoscopically on a large screen, there often exists a difference between the calculated depth and the perceived depth of an object. This paper presents a method for reducing such differences of depth. This is performed by modifying both the viewing position and the screen position in the stereoscopic calculation. The optimal amount of modification was decided using sample values of depth differences. Effectiveness of the proposed method is discussed upon experimental results. This technique decreased the average difference from 4.3mm to 1.3mm.

1 Introduction

Stereoscopic display is a necessary technique to observe 3D structures, to manipulate these structures directly, and to work intuitively in 3D virtual worlds. To make a 3D environment a creative workspace, a harmonized fusion of virtual and real worlds becomes necessary. To display correct stereoscopic effects, a pair of image, one for each eye, must be generated faithfully, based on both locations of the observer's eyes and of the screen [2]. However, some problems in stereoscopic display have been reported.

Fatigue of an observer and diplopia (double vision) are physiological problems caused by unnatural depth cues in stereoscopic display, such as the discrepancy between convergence and accommodation. To avoid these problems, the technique to render assuming a pupil distance smaller than the true distance as a parameter of stereoscopic images [13] is useful. However, this false pupil distance distorts the pair of images, and, as a result, makes the observer perceive incorrect depth [14].

Additionally, accuracy of viewing is another serious problem when doing some works in virtual environment. For example, to design a new digital tool such as computer aided design system with 3D canvas [12], virtual objects have to be observed in correct positions in real space, and a 3D virtual cursor must point to the same position as the physical pointer, which may be a pen-type sensor. For minimizing the difference between the modeled and the perceived world, accurate measurement of the parameters for creating stereoscopic images is necessary. However, to measure these parameters without error is too difficult in practice, and too faithful image depth such as ignoring the limits of fusion causes fatigue and diplopia depending on the situation [6].

These two important contrasting problems in stereoscopic display should be solved suitably depending on the needs of the application. If you desire only to feel 3D effects in virtual world for applying to visualization, such as landscape or flight simulator, the former may be important but high precision at closed-range is not needed. However, when you want to use it in industrial design field, the latter problem becomes important. This is because accuracy is a requirement in product design, so mismatches between the input device coordinates (real world) and the stereoscopically displayed coordinates (virtual world) must be avoided. The impression of virtual product made in un-harmonized coordinates will be different from the produced real product. This means that the examination and evaluation, which is the final stage of the design process, cannot be done correctly in 3D.

In this paper, we focus on the accuracy issues of the latter problem in stereoscopic display. In particular we discuss a method to improve accuracy using a large-sized screen displaying objects in front of the projection plane. Here, we intend that the method is useful for the design system using 3D space as a workspace.

1.1 The accuracy of depth perception

The first purpose in this field of study is to analyze the phenomenon of depth differences. The problem is complex because it relates not only to issues of computer graphics, such as tracking error, but also to the psychology of vision. Studies in both fields [8, 3, 5, 9, 7] have been extensively developed, but results so far have not been perfectly conclusive. Our interest here is directed to the modification of rendering parameters, inquiring how effectively it alone reduces depth differences and how reasonable it is.

For correction of depth perception by modification of parameters in stereoscopic rendering, pupil distance is often selected as the parameter because it changes depending on eyeball rotation [10]. If requirements on accuracy were not so serious, it would be a simple and useful method. But, this method works correctly only when the gazing point of an observer is already known, yet there are also other parameters bringing errors [11].

In accordance with the proposed method, firstly, parameters for modification are allocated based on geometrical relationship in image rendering. Next, the values of the parameters are determined to minimize sampled depth differences. The samples are perceived points that observers indicate using a physical pointing device. So if you want to measure only pure perceived depth, this procedure is not suitable for the measurement, because this is a simple comparison between the real object and the virtual one. However, we are able to regard this way as a fair measurement at least from the following two advantages. 1) The precision of measuring by device is better than the method of oral pointing proposed in previous works [4]. 2) Precision is usually necessary when the application deals with objects appearing near the observer. In this case users would manipulate virtual objects using a physical device.

In section 2, we allocate parameters for correction that translate the eye-points and the screen location in the stereoscopic rendering. We further present a procedure to determine the optimal modified values of these parameters from samples. In section 3, effectiveness of the proposed method is discussed in several aspects and experimental results are presented.

2 Registration Procedures

2.1 Parameter allocation

Figure 1 illustrates the parameters' relationship between the eye-points and the screen location for rendering of stereoscopic images such as head-tracked display systems [1, 12]. Gray objects denote positions before modification, which are usually measured by sensors and often present some errors. Black objects are positions modified by our proposed technique.

The constant p identifies the distance between the eye-points (pupil distance) and the variable s shows the distance between the middle point of the eyes and the screen (eye-and-screen distance) before registration. Parameters dp, ds_1 and ds_2 modify these distances to optimal ones $p + dp$ and $s + ds_1 - ds_2$. Both of ds_1 and ds_2 modify eye-and-screen distance, but ds_1 translates the eye-points and ds_2 translates the screen.

More parameters were not allocated because they do not influence depth differences significantly. For example, the distortion in image projection was calibrated sufficiently using a function of the display projector.

2.2 Optimal values of parameters

Let us suppose that, firstly, stereoscopic images are rendered with the eye-points and the screen location before correction (gray objects in Figure 1) and, secondly, they are observed with the same parameters after correction (black objects in Figure 1). Then the variable *y* denotes the theoretical value of the depth difference depending on variable *x*, which denotes the distance between the object and the screen (object-and-screen distance).
Thus, the following relationships are obtained from Figure 1.

$$\frac{p}{q} = \frac{s - x}{x} \tag{1}$$

$$\frac{p + dp}{q} = \frac{ds_1 + s - x - y}{x + y - ds_2} \tag{2}$$

When the variable *q* is eliminated from equation (1) and equation (2), they are arranged into the equation (3).

$$(s - x)(p + dp)(x + y - ds_2) = p\ x(ds_1 + s - x - y) \tag{3}$$

This equation is, then, rearranged as follows.

$$y = \frac{p\ x(ds_1 + s - x) - (s - x)(p + dp)(x - ds_2)}{(s - x)(p + dp) + p\ x} \tag{4}$$

After the pupil distance *p* is measured, we sample many sets of (*x*, *s*, *y*), and compute optimal values of (dp, ds_1, ds_2) as follows. When we define i^{th} sample as (x_i, s_i, y_i), the values of (dp, ds_1, ds_2) for each sample (x_i, s_i, y_i) exist and satisfy equation (4). These are computed for multiple samples. The optimal solution is obtained by minimizing the following function *g*.

$$f_i(dp, ds_1, ds_2) = y_i - \frac{p\ x_i(ds_1 + s_i - x_i) - (s_i - x_i)(p + dp)(x_i - ds_2)}{(s_i - x_i)(p + dp) + p\ x_i} \tag{5}$$

$$g(dp, ds_1, ds_2) = \sum_i \{ f_i(dp, ds_1, ds_2)\}^2 \tag{6}$$

At its minimum the $g(dp, ds_1, ds_2)$ satisfies:

$$\frac{\partial g}{\partial dp} = \frac{\partial g}{\partial ds_1} = \frac{\partial g}{\partial ds_2} = 0 \qquad (7)$$

This equation is solved numerically.

2.3 Practical solution

To numerically obtain a stable solution to equation (7) is too difficult because it is a non-linear minimization with three variables. Therefore we reduce the process to a uni-variable search as follows. We choose here to assign a certain small number to *dp* because we know that the margin of measurement error of the pupil distance is not large. Then, in order to minimize function (6), ds_1 and ds_2 can be obtained analytically.

When we define the following four variables:

$$\alpha_i = -p\ x_i / \delta_i \qquad (8)$$

$$\beta_i = -(s_i - x_i)(p + dp) / \delta_i \qquad (9)$$

$$\gamma_i = \{(s_i - x_i)(x_i + y_i)dp + p\ s_i y_i)\} / \delta_i \qquad (10)$$

$$\delta_i = (s_i - x_i)(p + dp) + p\ x_i \qquad (11)$$

the following function corresponds to function (5).

$$f_i(ds_1, ds_2) = \alpha_i ds_1 + \beta_i ds_2 + \gamma_i \qquad (12)$$

Then the following equations can be obtained by solving equation (7):

$$\begin{cases} \sum_i \alpha_i^2 ds_1 + \sum_i \alpha_i \beta_i ds_2 + \sum_i \alpha_i \gamma_i = 0 \\ \sum_i \alpha_i \beta_i ds_1 + \sum_i \beta_i^2 ds_2 + \sum_i \beta_i \gamma_i = 0 \end{cases} \qquad (13)$$

Next, we define the following five variables:

$$a_1 = \sum_i \alpha_i \gamma_i \qquad (14)$$

$$a_2 = \sum_i \alpha_i^2 \qquad (15)$$

$$b_1 = \sum_i \beta_i \gamma_i \tag{16}$$

$$b_2 = \sum_i \beta_i^2 \tag{17}$$

$$c = \sum_i \alpha_i \beta_i \tag{18}$$

ds_1 and ds_2 can be expressed:

$$ds_1 = \frac{b_1 c - a_1 b_2}{a_2 b_2 - c^2} \tag{19}$$

$$ds_2 = \frac{a_1 c - a_2 b_1}{a_2 b_2 - c^2} \tag{20}$$

A series of triplets (dp, ds_1, ds_2) is calculated for values of dp shifted around 0. The one minimizing the function (6) is selected as optimal.

3 Experiments

3.1 Display system

Figure 2 illustrates the display system. A pair of stereoscopic images (resolution of 1024×768 pixels) is projected on the screen (size of 2400mm×1800mm) from its rear side. We chose a large flat screen because a small screen that cannot cover the field of view has a harmful influence on depth perception, and a round screen is unsuitable for real-time rendering. Liquid crystal shuttered glasses are synchronized with the display for alternate projection to each eye. A viewing volume is defined by respective eye-points and the four corners of the screen. Objects are accurately rendered within this volume, both in position and size [1].

A cone (height 100mm, radius 25mm) is displayed to a subject and he points at the top with a stick. An optical sensor with high precision is used for position sensing.

The virtual object is, as closely as possible, colored and shaded similarly to a physical object. The whole display system is installed in a darkroom, and a checkered pattern is drawn on the screen as a background.

3.2 Results

Samples of depth differences were obtained from three subjects while object-and-screen distance and eye-and-screen distance were changed (Fig.3). The average value of five samples was calculated as a representative for each combination of four values of object-and-screen distance, 700mm, 800mm, 900mm and 1000mm and five values of eye-and-object distance, 300mm, 350mm, 400mm, 450mm and 500mm. That is:

s = 700mm, x = 200mm ~ 400mm

s = 800mm, x = 300mm ~ 500mm

s = 900mm, x = 400mm ~ 600mm

s = 1000mm, x = 500mm ~ 700mm

The optimal values of parameters for each person have been determined using the procedure described in section 2.

Red lines in the graphs show measured values, and light blue ones show the theoretical values of depth differences calculated with equation (4). Differences of y coordinates between red and light blue lines are the expected depth differences after registration. Their root-mean-square error values (RMS) are also shown, as a quantitative evaluation of correction.

3.3 Necessity of parameters

In Figure 3, sample distributions appear to vary among persons, but the differing values of dp enable us to explain them consistently. Figure 4 shows the registration results without ds_2. It indicates the necessity of ds_2 because these results do not fit the sample distributions and the values of dp seem unreasonable. When ds_1 or dp is ignored, similar results to those of ds_2 are obtained. Table 1 shows those results. Therefore, the cooperation of the three parameters is indispensable for effective registration.

3.4 Propriety of the values of parameters

All obtained values of *dp* in Figure 3 are several millimeters at the most. This is reasonable because pupil distances are measured with a caliper.

The appearance of ds_1 can be also explained by assuming that relative positions of eyes with respect to the screen contain errors.

However, non-zero values of ds_2 suggest reasons beside errors in position sensing. When the object is located on the screen, the observer should point just on the screen, because there is no disparity between stereo images on the screen, but non-zero values generate a disparity. The reasons are not clear yet. The appearance of ds_2 may be explained by the unnatural depth cues in stereoscopic display, such as the discrepancy between convergence and accommodation, or by assuming an individuals' habit in pointing action. When the observer always points a smaller depth than the correct one by ds_2, the eye-points' error in depth is equal to $ds_1 - ds_2$.

3.5 Verification

Finally, we verified the effect of our method with four subjects. Results of this verification are shown in Figure 5.

Groups of (x, s, y) were sampled for each combination of three eye-and-screen distances (s = 700mm, 800mm and 900mm) and three eye-and-object distances ($s - x$ = 300mm, 400mm and 500mm) (Fig.5a). The optimal values of registration parameters were fixed according to the procedures in section 2. Samples after registration were measured by using modified stereoscopic images that reflect the registration parameters (Fig.5b). Then, the depth difference y was measured for each combination of three eye-and-screen distances (s = 700mm, 800mm and 900mm) and five eye-and-object distances ($s - x$ = 300mm, 350mm, 400mm, 450mm and 500mm). In this figure, RMS' error is the root-mean-square of depth differences.

These results show that RMS' error decreases to 1.3mm in the average.

4 Conclusions

In this paper, we proposed a registration method of depth differences and discussed its adequacy and effectiveness. The method improves on our previously proposed method [11] in which the amount of correction of depth differences depended on both the object position and the screen position. More reliable values of registration parameters will be obtained from greater number of samples.

We applied this method to our Spatial Sketching System [12], but it may be applied to other head-tracking display environments, too. We further expect to develop an easier method to sample depth differences.

5 Acknowledgements

The authors would like to express their sincere gratitude to Associate Professor Geoff Wyvill for his encouragement and helpful advice for the accomplishment of this work. They also thank Dean and Professor Teruo Fukumura for his valuable comments and Christopher Houser and Pilar Almeida for their assistance in refining this paper. This work was supported in part by the grant-in-aid for Private University High-Tech Research Center from the Ministry of Education, Japan.

References

1. Cruz-Neira C, Sandin DJ, Defanti TA (1993) Surround-screen projection-based virtual reality: The design and implementation of the CAVE. Computer Graphics Proceedings, Annual Conference Series, 1993, 135-142
2. Deering M (1992) High Resolution Virtual Reality. COMPUTER GRAPHICS, Vol 26, No.2, 195-202
3. Kono H (1995) Development of virtual telephone system. International Conference on Artificial Reality and Tele-Existence / Conference on Virtual Reality Software and Technology '95, 39-46
4. Kurokawa T, Murata H, Miyanami M, Ogata M, Morimoto K (1996) Characteristics of depth perception in binocular stereoscopic vision using liquid-crystal glasses and its model. Progress in Human Interface, 5:61-68 (in Japanese)
5. Lampton DR, McDonald DP, Singer M (1995) Distance estimation in virtual environments. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting-1995, 1268-1272
6. Nagata S (1996) The binocular fusion of human vision on stereoscopic displays - field of view and environment effects. Ergonomics, Vol.39, No.11, 1273-1284
7. Ohtsuka S, Ishigure Y, Kanatsugu Y, Yoshida T, Usui S (1996) Virtual Window: A technique for correcting depth-perception distortion in stereoscopic displays. The Society for Information Display Symposium Digest, 27:893-898
8. Shibasaki H, Inoda K (1992) Study of 3D direct manipulation methods for CAD. 8th Symposium on Human Interface, 1-6 (in Japanese)
9. Shimizu H, Sunagawa S, Oda S, Miyachika K (1995) Study on 3-dimensional perception in stereoscopic display. Progress in Human Interface, 4:85-90 (in Japanese)
10. Utsumi A, Milgram P, Takemura H, Kishino F (1994) Investigation of errors in perception of stereoscopically presented virtual object locations in real display space. In proceedings of human factors and ergonomics society 38th annual meeting, 324-328
11. Yoshida S, Hoshino T, Ozeki T, Miyazaki S, Hasegawa J, Fukumura T (1998) An improvement method of depth perception in stereoscopic display. International Archives of Photogrammetry and Remote Sensing, Vol.XXXII, Part 5, 386-391
12. Yoshida S, Hoshino T, Ozeki T, Miyazaki S, Hasegawa J, Yasuda T, Yokoi S (2000) Spatial Sketch System for Car Styling Design, International Archives of Photogrammetry and Remote Sensing, Vol.XXXIII, Part B5, 919-925
13. Ware C, Gobrecht C, Paton MA (1998) Dynamic Adjustment of Stereo Display Parameters, IEEE Transactions on Systems, Man and Cybernetics, Vol 28. No 1, 56-65
14. Wartell Z, Hodges LF, Ribarsky W (1999) Balancing Fusion, Image Depth and Distortion in Stereoscopic Head-tracked Displays, SIGGRAPH 99, Computer Graphics Proceedings, Annual Conference Series, 351-358

Authors' biographies

SHUNSUKE YOSHIDA

Shunsuke Yoshida is a Ph.D. candidate at the Graduate School of Human Informatics, Nagoya University, Japan. He received his B.S. and M.S. in information science from Chukyo University in 1996 and 1998. His work focuses on the application of virtual reality to industrial design for practical use. His research interests include human interface, interactive application, computer-aided design, and virtual reality.

SHINYA MIYAZAKI

Shinya Miyazaki is an Associate Professor in the School of Computer and Cognitive Sciences at Chukyo University, Japan. He received his B.E., M.E., and Ph.D. degrees in information engineering from Nagoya University in 1989, 1991, and 1996. His research interests include computer graphics and virtual reality.

TOSHIHITO HOSHINO

Toshihito Hoshino is an assistant manager in the system-engineering department of the design division in Toyota Motor Corporation, Japan. He received his B.E. and M.E. in information engineering form Nagoya University in 1989 and 1991. His research interests include virtual reality, surface modeling, computer graphics, and animation.

TORU OZEKI

Toru Ozeki is an assistant manager in the system-engineering department of the design division in Toyota Motor Corporation.
He received his B.E. and M.E. in imaging science and engineering form Chiba University in 1983 and 1985. His research interests include virtual reality, color science, computer graphics, and animation.

JUNICHI HASEGAWA

Junichi Hasegawa is a Professor at the School of Computer and Cognitive Sciences at Chukyo University, Japan. He received a B.S. degree in electrical engineering, and M.S. and Ph.D. degrees in information science from Nagoya University in 1974, 1976 and 1979, respectively. He worked as a research associate and a lecturer of information engineering at Nagoya University from 1979 to 1987. His current research interests include image processing, pattern recognition, computer vision, and their applications to medical and sports science fields.

TAKAMI YASUDA

Takami Yasuda is an Associate Professor in the School of Informatics and Sciences at Nagoya University, Japan. He received his B.E. and M.E. in electronic engineering from Mie University in 1982 and 1984. He received his Ph.D. in information engineering from Nagoya University in 1989. His research interests include computer graphics and virtual reality on the Net.

SHIGEKI YOKOI

Shigeki Yokoi is a Professor in the Graduate School of Human Informatics at Nagoya University, Japan. He received his Doctoral degree in electrical engineering from Nagoya University. He has been an associate professor at Mie University and an associate professor at the School of Engineering at Nagoya University. His research interests include computer graphics, virtual reality, interactive media, and the social influence of digital technologies.

Table 1. Necessity of parameters

		all	without *dp*	without ds_1	without ds_2
Subject A	*dp*	2.08	0	-1.68	3.97
	ds_1	18.47	9.75	0	23.60
	ds_2	-6.56	-10.62	-11.42	0
	RMS	1.05	1.31	1.86	1.41
Subject B	*dp*	-1.80	0	-3.69	1.04
	ds_1	9.37	16.91	0	17.04
	ds_2	-10.76	-6.99	-13.47	0
	RMS	1.22	1.40	1.46	1.91
Subject C	*dp*	-1.83	0	-1.31	6.19
	ds_1	-2.66	5.40	0	19.44
	ds_2	-29.95	-25.99	-29.17	0
	RMS	1.28	1.47	1.30	4.46

(unit:mm)

list of figure legends

Fig.1. Arrangement of parameters for modification

Parameters *p*, *s*, *x* indicate the pupil distance before correction, the eye-and-screen distance and the object-and-screen distance, respectively. *dp* is the modified value for pupil distance; ds_1 and ds_2 are modified values for the eye-and-screen distance. *q* is the difference between displayed objects on stereo images. Finally, *y* indicates the corrected value for the object's depth.

Fig.2. Structure of experimental system

Fig.3. Relationship between object-and-screen distance and depth difference

Red bend lines in the graphs show measured values, and light blue ones show the theoretical values of depth differences calculated with equation (4). RMS is the root-mean-square of differences of *y* coordinates between corresponding nodes on red bend lines and light blue lines.

Fig.4. Results without ds_2 (compare with Fig.3.)

Red bend lines in the graphs show measured values, and light blue lines show the theoretical values of depth differences calculated with equation (4). RMS is the root-mean-square of differences of *y* coordinates between corresponding nodes on red bend lines and light blue lines.

Fig.5. Result before and after correction

Red bend lines in the graphs show measured values, and light blue lines show the theoretical values of depth differences calculated with equation (4). RMS' means root-mean-square of depth differences.