Depth cue combinations for density judgment in three-dimensional display

Kentaro Yamamoto[*†], Ricky K. C. Au[†‡] & Katsumi Watanabe[*†]
[*] Waseda University, Tokyo, Japan
[†]The University of Tokyo, Tokyo, Japan
[‡]The Hong Kong Polytechnic University, Hong Kong, China
E-mail: yamamoto.kntr@gmail.com Tel: +81-03-3209-3211 (3392)
E-mail: rickyau10@gmail.com Tel: +852-2766-7094
E-mail: katz@waseda.jp Tel: +81-03-3209-3211 (3391)

Abstract—**The present study examined whether different depth cues would interact with the judgment of density differently. Using a 3D projector, we presented random-dot stimuli on fronto-parallel square planes at different depths (-30, -15, 0, +15, or +30 cm from the projection plane) and measured the perceived density of the dots in each depth plane using the method of constant stimuli. The depth of the plane was manipulated with three types of depth cues (binocular disparity, stimulus area, and dot size), which were used separately or all together. The results showed that the averaged PSEs depended on the depth plane when the depth cue was stimulus area, whereas the influences of size and disparity cues on the PSE were relatively weak. However, when the cues were combined, the influence of the area cue on density judgment was largely attenuated. These findings imply that the combination of the depth cues can provide more precise depth perception of the dots and helping interpretations of "3D-valid density".**

I. Introduction

With the recent advances in science and technology, there is increasing interest in three-dimensional (3D) information for the purpose of sensing and visualization of real-world objects, and 3D technology has been widely used in various fields including entertainment, business, and medical services [1]. Stereoscopic vision provides richer information about distance and depth perception, which enhances the realistic sensation of virtual scenes and enables more precise remote operations. Recently, binocular disparity cue, which indicates the difference in position of objects seen by the left and the right eyes, has become popular for producing stereoscopic vision in 3D devices. However, humans perceive depth using many other different cues such as retinal image size [2], perspective [3], and shadow [4]. Therefore, for more effective use of 3D information, it is important to investigate how the different cues interact with each other to provide depth perception.

Previous studies have examined the effect of the combination of depth cues on perceived depth and proposed a linear model and a multiplicative model [5]. In the linear model, depth cues are supposed to be computed independently and combined linearly (additively). On the other hand, in the multiplicative model, the computation of depth is not clearly separated among depth cues and the cues are supposed to interact with each other. Although these models expect completely different results, previous studies have found that the cue combination might be based on both the linear and the multiplicative models depending on the experiment. Based on the discrepancy, Landy et al. [5] proposed an alternative depth cue combination model, where

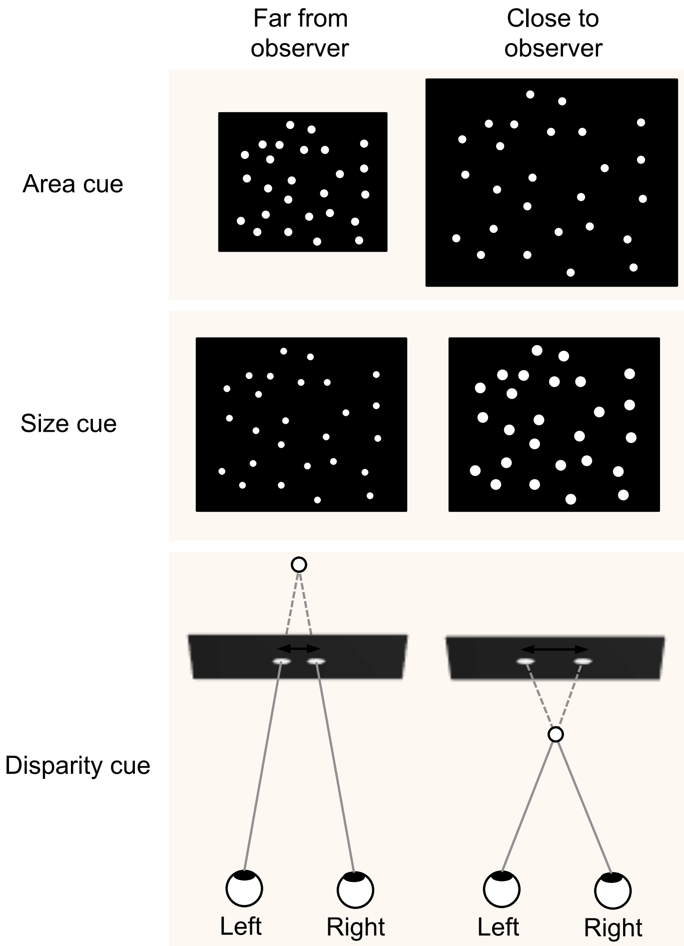

Fig. 1. Three types of depth cues manipulated in this study. Area cue indicates the total area occupied by the dots. Size cue indicates the projected (retinal) size of the dots. Disparity cue indicates the difference in position of dots seen by the left and the right eyes.

depth cues are first computed independently, and then interact to be commensurate with each other and combined linearly (in small perturbation case) at the end. Although this model can well-explain the results of the previous studies, most of the studies have measured perceived depth directly using localization and depth comparison tasks and it was unclear how depth cues interact with each other when people judge perceived depth indirectly.

In the present study, we manipulated the presentation depth of random-dot stimuli with three types of depth cues and examined the separate and combined effects of the different cues on density judgment. Density is defined as mass per unit area or volume. Particularly, in the case of 2D random-dot patterns, density is generally defined as the number of dots per area occupied by the dots [6]. This means that density values are reduced as the total stimulus area becomes larger. Although retinal size of stimulus area changes as a function of observation distance, its perceptual size is kept almost constant over the distance. This is because our perception of size is based not only on retinal image size but also on distance, enabling us to make accurate estimates of the real size (size constancy [7]). Thus, distance and depth perception are both important for size estimation and, in turn, for density judgment. Based on this, we used density judgment as a measure of depth perception. We expected that density judgment would be influenced by the stimulus area, which would change depending on the projected area and the perceived depth.

In the present experiment, the depth cues, "Area", "Size", and "Disparity", were manipulated separately. Fig. 1 shows schematic illustrations of each manipulation. With the Area cue, the total projected area is larger as the plane of dots is closer to the observer. With the Size cue, the projected dot size is larger as the plane of dots is closer. With the Disparity cue, the disparity between the left- and the right-eye images changes according to the depth of the plane of dots. There were six cue conditions in the present experiment: Area cue only, Size cue only, Disparity cue only, Area and Size cues combined, Area and Disparity cues combined, and All cues combined. Each observer performed all the cue conditions.

II. Method

A. Observers

Two psychophysically experienced observers (KY and KT) participated in the experiment. Both of them had normal or corrected-to-normal visual acuity, and provided informed consent prior to the study.

B. Apparatus

The stimuli were programmed in MATLAB R2012b (MathWorks, USA) using the Psychophysics Toolbox extension (version 3.0.8 [8, 9]). The 3D stimuli were produced using the OpenGL functions embedded in the toolbox. The stimuli were projected by a 3D projector (Sight3D, Solidray Co. Ltd., Japan) at a refresh rate of 60 Hz, and were displayed on a flat surface (170 cm × 127.5 cm).

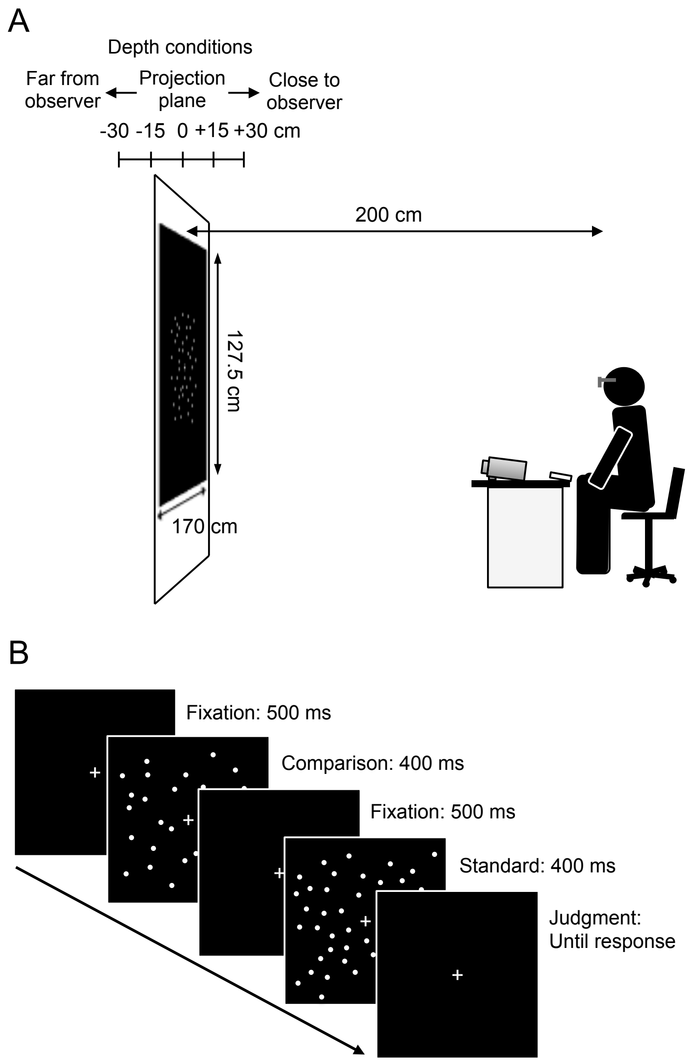

Fig. 2. Schematic illustrations of the apparatus (A) and the procedure (B).

The projector was controlled by a personal computer running the Windows 7 operating system. The observers seated at a distance of 200 cm away from the projection screen and viewed the projected stimuli using 3D glasses (Fig. 2A).

C. Stimuli and Procedure

In each trial, two random dot stimuli (the standard stimulus and the comparison stimulus) were displayed sequentially in counter-balanced order. Fig. 2B shows a schematic illustration of a single trial. Each trial began with a blank screen, and the observer initiated the trial by pressing the space bar on the keyboard. A fixation cross (0.6 deg × 0.6 deg) appeared at the center of the screen for 500 ms. Then, the first stimulus appeared for 400 ms and followed by another fixation screen for 500 ms. The second stimulus then appeared for 400 ms, followed by another fixation screen. The observer was asked to determine whether the first or the second stimulus had a higher density of dots. The observer made the judgment by pressing the predetermined keys. Then, the trial ended with the disappearance of the fixation cross.

The dots (virtual spheres of 0.6 cm in diameter) in each stimulus were presented on a fronto-parallel square plane with a size of x (horizontal) × y (vertical) = 60.5 cm × 60.5 cm at the distance of 200 cm. The dots were presented in random x- and y- positions while avoiding overlapping. The standard stimulus was always at the projection plane with 50 dots. The comparison stimulus contained a total of 10, 20, 30, 40, 50, 60, 70, 80, or 90 dots and at a fronto-parallel plane of either -30, -15, 0, +15, or +30 cm from the projection plane. For example, in the +30 cm condition, the dots were at the plane of 30 cm closer to the observer than the projection plane, i.e., at a distance of 170 cm from the observer.

Depth planes of the comparison stimulus were manipulated by three depth cues: (a) the total area occupied by the dots projected on the projection plane depending on the simulated depth with perspective (52.6 cm × 52.6 cm in the farthest (-30 cm) condition, 60.5 cm × 60.5cm in the no depth (0 cm) condition, and 71.2cm × 71.2cm in the closest (+30 cm) condition), (b) the size of dots projected on the projection plane (0.52 cm × 0.52 cm in the farthest (-30 cm) condition, 0.60 cm × 0.60 cm in the no depth (0 cm) condition, and 0.71 cm × 0.71 cm in the closest (+30 cm) condition), and (c) the binocular disparity corresponding with depth (0.85 cm in the farthest (-30 cm) condition, 0 cm in the no depth (0 cm) condition, and -1.15 cm in the closest (+30 cm) condition).

The experiment consisted of a total of 30 conditions: 6 Cue conditions (Area, Size, Disparity, Area+Size, Area+Disparity, All) × 5 depth conditions of the comparison stimulus (-30, -15, 0, +15, +30 cm). The cue was fixed within a session, and each session consisted of 270 trials: 5 depth × 9 dot-numbers × 2 presentation orders × 3 repetitions. Each observer repeated twice for each session, resulting in 12 sessions (3240 trials) in total.

III. Results

The proportion of trials that the comparison stimulus was judged to have higher density than the standard stimulus was calculated and plotted as a function of the number of dots in the comparison stimulus. Then, the data of each observer in each depth and cue condition were separately fitted with a logistic function (1):

$$f(x) = 1/(1 + \exp[\beta(x_0 - x)]) \qquad (1)$$

where x_0 is the x-value at the midpoint, and β is slope of the curve. Fig. 3 shows an example of psychometric function from a typical condition. The point of subjective equality (PSE), at which the number of dots in the comparison stimulus that the observers judged to be equal in density to that of the standard stimulus (i.e. midpoint of the psychometric function), was computed and used for the analysis.

Fig. 4A shows the averaged PSEs of the two observers in each condition. A 6 (depth cue) × 5 (comparison depth) two-way within-subject analysis of variance (ANOVA) showed neither a significant main effect of depth cue ($F(5, 5) = 1.62$, $p = .30$) nor a significant main effect of comparison depth ($F(4, 20) = 4.31$, $p = .09$). However, there was a significant interaction between depth cue and comparison depth ($F(20, 20) = 5.12$, $p < .001$). Simple main effect analyses based on the interaction revealed significant main effect of comparison depth in the Area cue and Disparity cue conditions (Area: $F(4, 24) = 21.13$, $p < .001$; Disparity: $F(4, 24) = 3.13$, $p < .05$), but not in the other cue conditions (Fs < 2.5, ps > .07).

The results of the PSE analyses suggest that the perceived density changes depending on both the comparison depth and the depth cue. Specifically, with the Area cue only, the PSE became higher (i.e., the perceived density was decreased) as the dots were presented closer to the observer. This indicates that the density of dots appeared sparser as the projected (retinal) size of the stimulus area was larger; and this is not surprising because an increased area with fixed dot number means an increased average space between the dots. On the other hand, the differences in PSE among the comparison depths were negligible in the other conditions except for the Disparity cue condition.

For a more direct comparison of the influence among the depth cue conditions, we also computed the linear slopes of PSE change as a function of the presentation depth of the comparison stimulus. As shown in Fig. 4B, the slope value in the Area cue condition was positive and relatively large in each observer, while the slope values in the Size and Disparity cue conditions were negative and small. When the cues were combined, however, the steep slope observed in the Area cue condition became flattened; the PSE slope values were closer to zero in the Area and Size cues, Area and Disparity cues, and All cues conditions. These differences were also confirmed by a one-way ANOVA which showed a significant main effect of depth cue ($F(1, 5) = 6.56$, $p < .05$).

These findings imply that, although the size and disparity cues per se have little effect on density perception, the

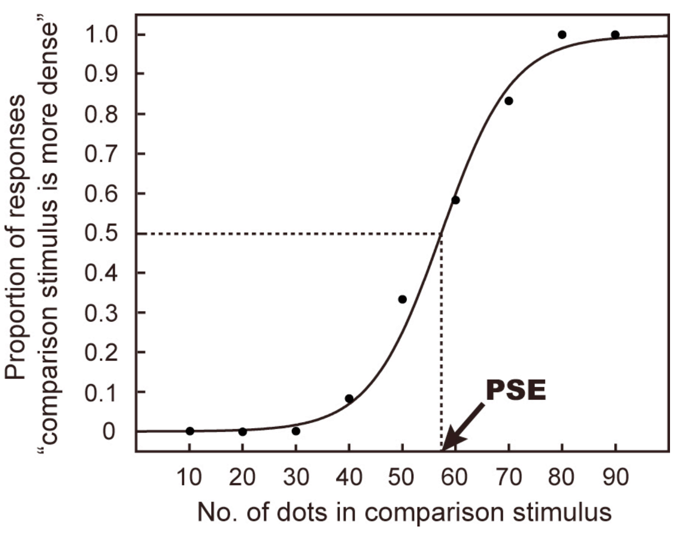

Fig. 3. Psychometric function obtained from one of the observers when the dots were presented at -15 cm with the disparity cue. The proportion of responses that the comparison stimulus had higher density than the standard stimulus was plotted against the number of dots in the comparison stimulus. The point of subjective equality (PSE) is indicated by the arrow.

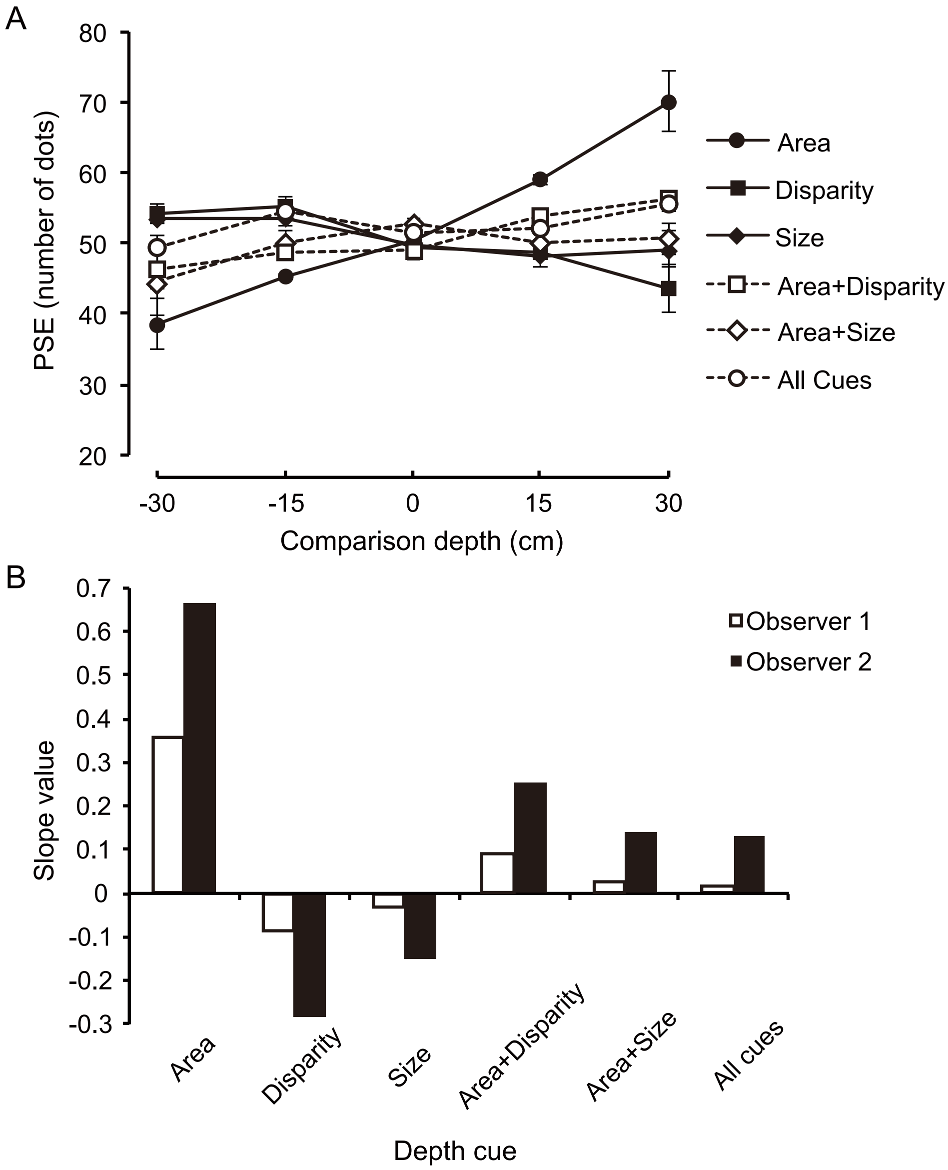

Fig. 4. Results of the experiment. (A) Averaged PSEs in each depth cue condition as a function of comparison depth. Error bars indicate the standard error. (B) Slope values of linear regression lines fitted to PSE change in each depth cue condition.

combination of the two or three depth cues reduces the effect of the area cue.

IV. Discussion

The present study examined the separate and combined effects of different depth cues on density judgment in the random-dot display. We found that the stimulus area strongly influenced density judgment, while the dot size and binocular disparity per se had less impact. However, if each manipulation was combined with the manipulation of the stimulus area, the strong influence of stimulus area on density judgment was largely attenuated. This attenuation effect was more pronounced when all cues were manipulated together. Given that the projected stimulus area is the same as the Area cue condition, the attenuation effect can be attributed to the change in perceived depth of stimulus presentation. These findings suggest that depth cue combination is important for providing more precise depth perception of the dots and helping interpretations of “3D valid density.”

When the disparity cue was manipulated alone, the perceived density significantly increased as the presentation plane was closer to the observer. This influence was opposite to the effect of the area cue. This is because, in the Disparity cue condition, the projected area was unchanged irrespective of the comparison depth, and thus the perceived stimulus area should be changed in the opposite direction from the Area cue condition due to size constancy [6]. In contrast, the influence of the size cue on density judgment was not significant. These results suggest that, different from pictorial (monocular) depth cues such as stimulus size and area, binocular depth cue is enough to influence density judgment by itself, though the magnitude of the effect was much weaker than when stimulus area was physically manipulated. This is consistent with previous studies that suggested that binocular disparity alone can be used as a depth cue [10].

Although the area and size cues were not sufficient to provide depth perception when they were manipulated independently, the erroneous density judgment due to stimulus area largely disappeared when the cues were combined, suggesting that area and size cues may be regarded as depth information only when they are manipulated together. This is inconsistent with the previous models, which predict linear or multiplicative effect of cue combination, and indicates the specificity of the area and size cues on depth perception. Because retinal size changes with actual size of the object and observation distance, it is possible that perceptual system interpret that actual size rather than observation distance is changed when only stimulus area or size is changed. The present results may suggest the importance of manipulating stimulus area and size together when they are used as depth cues.

V. Conclusions

Here we present evidence that the combination of different depth cues is important to provide enough depth information to judge the density of random-dot patterns, which were presented in different depth planes, without being affected by the retinal size of stimulus area. In addition, the disparity cue influenced density judgment by itself, indicating that binocular disparity alone is sufficient for stereoscopic depth perception. Our findings suggest the importance of combining different kinds of depth cues for more effective application of 3D information.

Acknowledgment

This research was supported in part by Grant-in-Aid for JSPS Fellows to KY, Grant-in-Aid for Scientific Research (23240034), and CREST, Japan Science and Technology Agency to KW.

References

[1] R. Bajcsy, R. Yang, P. Zanuttigh, and C. Zhang, “3D Imaging Techniques and Multimedia Applications,” *IEEE MultiMedia*, vol. 20, pp. 14–16, 2013.
[2] D. Regan and K. I. Beverley, “Binocular and monocular stimuli for motion in depth: changing-disparity and changing-size feed the same motion-in-depth stage,” *Vision Res*., Vol. 19, pp. 1331–1342, 1979.
[3] M. L. Braunstein, “Depth perception in rotating dot patterns, Effects of numerosity and perspective,” *J. Exp. Psychol*., vol. 64, pp. 415–420, 1962.
[4] D. Kersten, D. C. Knill, P. M. Mamassian, and I. Bulthoff, “Illusory Motion from Shadows,” *Nature,* vol. 379, pp. 31, 1996.
[5] M. S. Landy, L. T. Maloney, E. B. Johnston, and M. Young, “Measurement and Modeling of Depth Cue Combination- in Defense of Weak Fusion,” *Vision Res*., Vol. 35, pp. 389–412, 1995.
[6] S. C. Dakin, M. S. Tibber, J. A. Greenwood, F. A. Kingdom, and M. J. Morgan, “A common visual metric for approximate number and density,” *Proc. Nati. Acad. Sci. USA.*, vol. 108(49), pp. 19552–19557, 2011.
[7] R. L. Gregory, *Eye and Brain*. New York: McGraw-Hill, 1966.
[8] D. H. Brainard, “The psychophysics toolbox,” *Spat. Vis.*, vol. 10(4), pp. 433–436, 1997.
[9] D. G. Pelli, “The VideoToolbox software for visual psychophysics: Transforming numbers into movies,” *Spat. Vis.*, vol. 10(4), pp. 437–442, 1997.
[10] B. Julez, *Foundations of Cyclopean Perception*. Chicago, IL: The University of Chicago Press, 1971.