Two independent mechanisms for perception of motion in depth

Satoshi Shioiri

Department of Information and Image Sciences, Chiba University,
Chiba City 263-8522, JAPAN
E-mail: shioiri@image.tp.chiba-u.ac.jp

Two different binocular cues are known for detecting motion in depth. One is disparity change in time and the other is inter-ocular velocity difference. Shioiri, Saisho and Yaguchi (1999) demonstrated that motion in depth can be seen based solely on inter-ocular velocity differences as well as on the disparity change in time. They used conditions in which either cue was minimized and measured performance based on motion in depth, finding better performance than chance level when either velocity cue or the disparity cue was almost isolated. However, there may have been influences from the cue minimized in each condition, since it was practically impossible to isolate perfectly either cue. I re-analyzed their data to examine whether the performance in the condition with disparity change and that in the condition with inter-ocular velocity difference were correlated. The result showed the correlation is very low and therefore, we can conclude that the visual system has two different mechanisms for motion in depth.

I. INTRODUCTION

One of the most important functions of the human visual system is to reconstruct the three-dimensional world from two-dimensional retinal images. There are a number of useful cues for the 3D perception. Many of them are found in paintings, and they include perspectives, shadow and shading, size change, interposition, and so on. We can enjoy TV, movies, and photographs because of these, the so-called pictorial cues. In addition to the process of seeing depth from these pictorial cues, the human visual system has a mechanism to see depth based on binocular retinal images, i.e., binocular stereopsis. Since the two eyes look at much the same region of visual space, the difference between the two retinal images can be used to calculate the distance in depth between the objects seen. Both binocular stereopsis and depth perception of pictorial cues, including the interaction among, cues have been investigated extensively for decades by visual scientists.

For moving objects, there are also multiple cues for 3D perception. Among a number of cues, binocular cues have been investigated in detail. There are three major facts that indicate that the human visual system has binocular mechanism(s) that specialize in detecting motion in depth. First, the detection threshold of motion in depth is elevated by selective adaptation of motion in depth stimuli, which produces little threshold change for detecting static disparity (Regan & Beverley, 1973; Beverley & Regan, 1973). This implies that there are independent mechanisms for detecting static depth and motion in depth. Second, threshold and suprathreshold responses to a pulse of disparity are different for approaching and receding pulses (Beverley & Regan, 1974). Such a difference between motion directions should be attributed to a motion sensitive mechanism. Third, the binocular visual field of normal observers contains areas with no sensation of motion in depth for disparity pulses whereas motion in depth is seen for the pulses in the opposite direction (Hong & Regan 1989).

There are two possible binocular mechanisms for two different binocular motion cues (Fig. 1). (Cumming & Parker, 1994; Regan 1993; Shioiri, 1995; Shioiri, Saisho & Yaguchi, 1999). One is that based on disparity change in time and the other is that based on inter-ocular velocity differences. In the former case, disparity is detected first, and then its change over time is calculated. In the latter case, monocular velocities are calculated first, and then compared. Since motion in depth can be seen in dynamic random-dot stereograms, where no coherent monocular motion exists, disparity change in time is used to see motion in depth (Julesz, 1971; Norcia & Tyler, 1984). In contrast, there is a disagreement among studies for the contribution of inter-ocular velocity differences to motion in depth (Cumming & Parker, 1994; Nagata, 1982; Shioiri, Saisho & Yaguchi, 1999). Cumming and Parker, for exam-

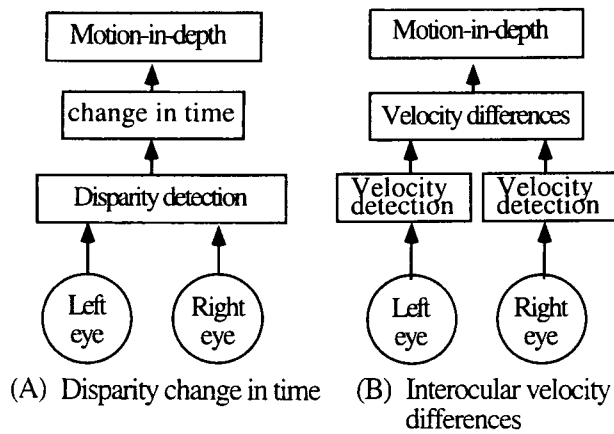(A) Disparity change in time   (B) Interocular velocity differences

FIG. 1. Two possible mechanisms for perception of motion in depth. (A) Disparity signal is detected and the perception of motion in depth is processed based on its change in time signals. (B) Velocity signal is detected in each retinal image and they are compared to provide the motion in depth perception.

FIG. 2. Binocularly uncorrelated random-dot kinematograms (top) and the motion signal in each eye (bottom).

ple, reported little contribution of inter-ocular velocity differences to motion in depth whereas Shioiri et al. (1999) showed clear contribution of inter-ocular velocity differences. In this report, I reanalyzed the data of the experiments of Shioiri et al. (1999) to support the idea that there are two different mechanisms to detect the two different cues to see motion in depth.

The variabilities in performance are the prime interest of this report. Shioiri et al. used two different stimuli: one was the stimulus to isolate the inter-ocular velocity differences (velocity cue condition) and the other was to isolate the disparity change in time (disparity cue condition). The differences of performance among different experimental conditions are expected to be uncorrelated between the two cue conditions if two different mechanisms detect the motion-in-depth signal in the two different cue conditions. On the other hand, a strong correlation is expected between the two conditions if a single mechanism detects the motion-in-depth signal in both conditions. This analysis is important since there are possible residual disparity cues in the velocity cue condition as there are possible residual velocity cues in the disparity condition. In this report, I first describe the experiments of Shioiri et al, and then, the analysis of the correlation of performance between the conditions.

II. MOTION IN DEPTH FROM INTER-OCULAR VELOCITY DIFFERENCES

To examine whether motion in depth can be seen solely based on inter-ocular velocity differences, Shioiri et al. (1999) used binocularly uncorrelated random-
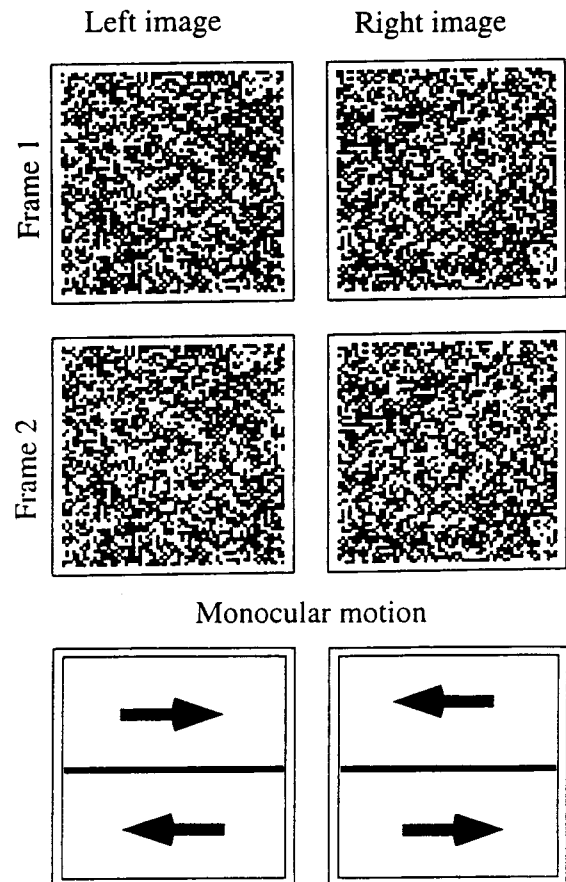
dot kinematograms (Fig. 2). The kinematograms contained two frames for each eye and the displacement of the dots of the left image was in the opposite direction to that of the right image. There was no correlation between the left and right images, and therefore, no binocular disparity cue was assumed to be available for depth processing (fusing the images in Fig.2 provide an unstable surface with lacy depth). In each random-dot kinematogram, the upper half of the pattern moved leftward and the lower half moved rightward or vice versa, creating relative motion between the halves. The replacements of the frames provided motion signals in opposite direction for the two retinae in each half of the images. The observers indicated the direction of motion in depth of the upper half of the stimulus (moved forward or away) and the percentages of correct responses were recorded for different contrast levels of the stimulus by the method of constant stimuli. Each random-dot field consisted of 256 × 256 pixels which corresponds to 4.3° × 4.3° in visual angle. The displacement size (the shift size of dots between the two frames) was either 4, 8 or 16 min of arc in visual angle.
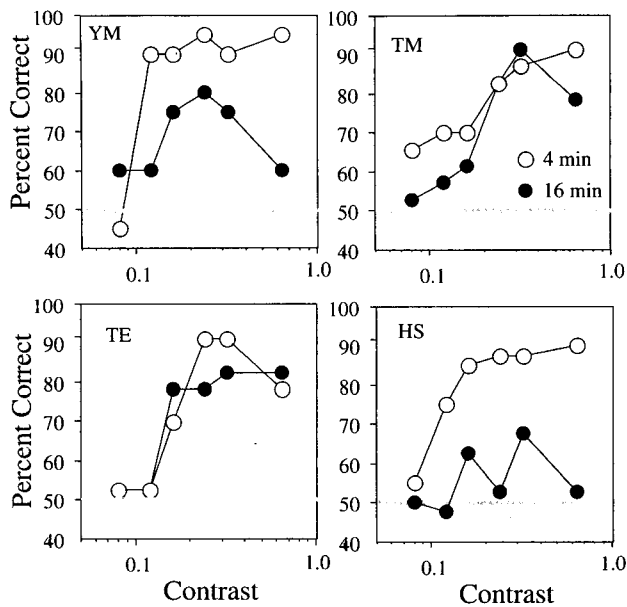
FIG. 3. Percentage of correct responses as a function of stimulus contrast for two displacement sizes. Each panel shows results of each observer.
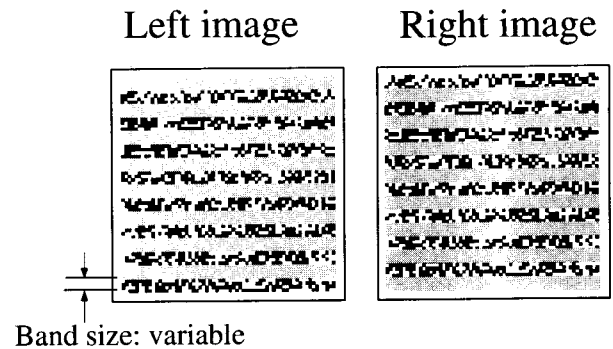

FIG. 4. Stimulus to avoid possible random correspondences. There is no direct overlap of the dots in left and right retinal images with appropriate binocular fusion.
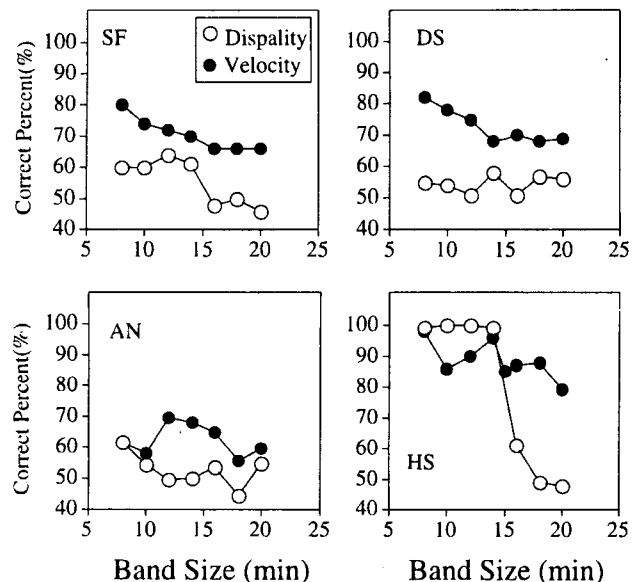

FIG. 5. Percentage of correct responses as a function of band size for the stimulus without direct overlaps in the velocity and disparity cue conditions.

Fig. 3 shows the average percentages of correct responses as a function of contrast for displacement size of 4 and 16 min, separately for four observers. Each datum point was based on 60 judgments. The line of 50% indicates the chance level of performance.

The results were similar for all the observers, showing that the performance was higher than chance level for contrasts of 0.2 or higher for all displacement sizes. The effect of displacement size was also similar among the observers. The highest percentage of correct responses was found for the smallest displacement size and the percentages decreased with displacement size (the results for 8 min displacement, which are not shown, are between the data of 4 and 16 min displacements).

III. EFFECT OF RANDOM CORRESPONDENCES

An alternative interpretation of the results in Fig. 3 is that randomly paired dots in the binocularly uncorrelated random-dot kinematograms provided a disparity signal that changed in time. Although there was no correlation between the left and right images, a certain number of dots could be paired with randomly varying disparity. One technique to minimize the effect of the residual disparity cues in uncorrelated random-dot kinematograms was proposed by Shioiri et al. In the stimulus of the second experiment of Shioiri et al., parts of the random-dot display in each eye were replaced with uniform gray horizontal bands

that alternated in vertical position in the left and right images (Fig. 4). Each image consisted of a square-wave horizontal grating of random-dot bands with gray bands. Since the bands with dots occupied alternate positions in the left and right images, there was no overlap with appropriate binocular fusion. They also used the temporally uncorrelated random-dot stereogram (also called dynamic random-dot stereogram) to isolate disparity cues for comparison. The correlation between the left and right eyes was 100% while the correlation between the first and the second frames was 0%. Monocular observation of the stimulus provided no coherent motion.

Fig. 5 shows the percentage of correct responses as a function of the band size for the results of the two cue conditions of one observers. The dot contrast was fixed at 0.32 in this experiment. The results showed

that the observer could identify the direction of motion in depth even without direct binocular overlaps of dots in both conditions. Performance was better in the velocity condition than in the disparity condition and the difference is larger with larger band sizes. This strongly suggests that inter-ocular velocity differences provide motion in depth signals in the visual system.

On the other hand, it is not clear from the data whether the motion detection in the disparity cue condition was based on the detection of disparity and its change in time. There should be random correspondences of dots across the first and second frames in either left or right images to provide motion signal in a direction. The other image should have a corresponding part that moves with a different velocity. The difference in the velocity between the left and right retinae may used to see motion in depth as in the velocity cue condition. This is the counterpart of the random correspondence problem in binocularly uncorrelated random-dot kinematograms.

IV. CORRELATION BETWEEN THE CONDITIONS

The fact that motion in depth can be seen both in the velocity and disparity cue conditions suggests that there are two different mechanisms for motion in depth perception. However, it is practically impossible to isolate perfectly each cue in the stimulus conditions as described. To examine whether these two cues are detected by a single mechanism or not, the correlation is taken between the percentages of correct responses in the two conditions. If there were only one mechanism to detect motion in depth, inter-condition differences should be similar in the two cue conditions. Consequently, the correlation of the performance between the two conditions would be high. In contrast, the correlation was expected to be much lower if different mechanisms detect motion in the two different cue conditions.

Fig. 6 shows the correlation of the percentages of correct responses for the identification of motion in depth between the velocity and disparity conditions in the second experiment of Shioiri *et al.*. Different symbols represent the results of different observers. The figure shows that the data points are rather scattered. The correlation is low: the average correlation coefficient across observers was 0.27 (−0.01, 0.74, −0.33, and 0.67 for four observers). Since the performance is expected to decline with band size for both conditions, there should be some correlation between the results. However, the small correlation coefficients suggest that the effect of band size is different for the use of the two cues.

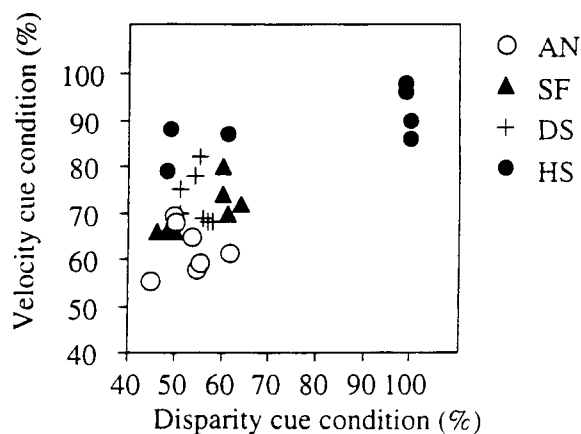The poor correlation between the two conditions

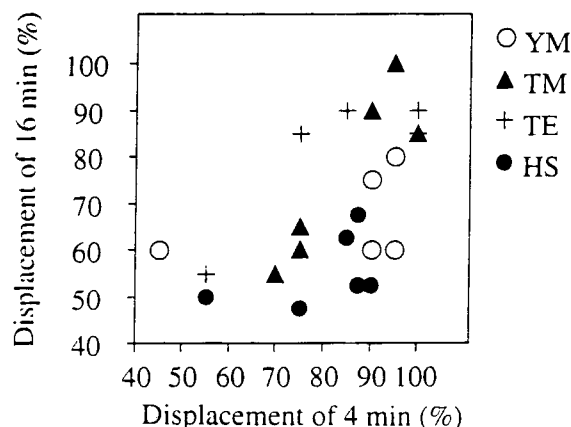FIG. 6. Correlation of percentage of correct responses between the velocity and disparity conditions.

FIG. 7. Correlation of percentage of correct responses between the 4 min and 16 min displacement conditions.

suggests that there are two different mechanisms to detect motion in depth in the velocity and disparity cue conditions. However, it has to be examined whether the correlation is higher when a single mechanism detects the motion signal in the two different conditions. To this purpose, I compare the data in the two displacement conditions in the first experiment, where the same mechanism was assumed to detect the motion in depth. Fig. 7 shows the correlation between the percentage of correct responses of the 4 min and that of the 16 min displacements. The data points are plotted around a straight line and the average correlation coefficient was 0.74 (0.54, 0.89, 0.98 and 0.55 for four observers). These coefficients are higher than the values between the different cue conditions. This supports the idea that the poor correlation between the cue conditions is due to the difference of mechanisms for the detection of the motion-in-depth signal in depth.

V. DISCUSSION

We reanalyzed the data of Shioiri et al. (1999) to examine whether there are two different mechanisms to detect the two different cues to see motion in depth. The results showed that the correlation coefficient of the performance between the velocity and disparity conditions is low whereas that between the different displacement conditions of the velocity condition is higher. This suggests that there are two different mechanisms to detect motion in depth, one based on disparity change in time and the other based on inter-ocular velocity differences, agreeing the conclusion of Shioiri et al..

Is it necessary to have two different mechanisms for detecting one feature? If the sensitivity of one of them is always less than that of the other, the role of the mechanism with lower sensitivity is perhaps trivial. The experiments by Cumming and Parker (1994) suggest that the disparity detection mechanism is more sensitive than the velocity detection mechanism over a wide range of spatio-temporal frequencies. They reported that the sensitivity for motion in depth in the condition where only velocity cues were available was too low to measure. On the other hand, in Fig. 5 higher sensitivity was shown in the velocity cue condition than in the disparity cue condition. The result provides an estimate of the integration limit of the information of left and right retinae. Since the performance dropped to chance level at about the band size of 15 min in the disparity condition(except for DS, whose performance was around chance level for all the band sizes), the integration limit is suggested to be about 15 min. In contrast, the decline of the performance was rather gradual and the performance never reached the chance level within the range of band size used. This suggests that the integration limit for the velocity cue is larger than that for the disparity cue and that inter-ocular velocity difference may be used dominantly for larger stimuli or stimuli that required large spatial integration.

There is another possible difference between the use of the two cues. Detecting disparity is likely a slow process while detecting motion is likely a fast process. If this is the case, we expect better sensitivity to velocity differences at faster speeds than to disparity changes and vice versa at slower speeds. Preliminary experiments in our laboratories showed results that support this. The performance for direction discrimination of motion in depth was better in the disparity cue condition at a slower speed while it was better in the velocity condition at a faster speed. It seems, therefore, that the mechanism to detect inter-ocular velocity differences works for perceiving motion of moving objects with higher speeds than that to detect disparity change in time. This issue, however, needs further investigation since the preliminary experiment used only three velocities and the difference of performance between the two cue conditions was not very large.

VI. CONCLUSION

The correlation analysis suggests that the visual system has two different mechanisms to see motion in depth, which are sensitive to different types of stimulation. These two mechanisms may be specialized for processing different types of moving objects.

ACKNOWLEDGMENTS

REFERENCES

[1] Beverley K. I. & Regan D., Journal of Physiology **235**, 17 (1973).
[2] Beverley K. I. & Regan D., Vision Research **14**, 357 (1974).
[3] Cumming, B. G. & Parker, A. J., Vision Research **34**, 483 (1994).
[4] Hong X. & Regan D., Vision Research **29**, 809 (1989).
[5] Julesz B., *Foundations of cyclopean perception.* (Univ. Press., Chicago, 1971).
[6] Nagata, S., NHK Science and Technology Research Laboratories Monthly Reports **25**, 487 (1982).
[7] Norcia A. M. & Tyler C. W., Vision Research **24**, 395 (1984).
[8] Regan D., Depth from motion and motion-in-depth. In *Vision and visual dysfunction Vol. 9, Binocular vision* (ed. Regan, D), London: MacMillan, 137 (1991).
[9] Regan D., Vision Research **33**, 2359 (1993).
[10] Regan D. & Beverley K. I., Science **181**, 877 (1973).
[11] Shioiri, S., Spatial frequency characteristics for binocular depth and motion in depth, Proceedings of 2nd International Display Workshops **3**, 61 (1995).
[12] Shioiri S, Saisho, H. and Yaguchi H., Vision Research, in submission (1999).