Japanese Psychological Research
THE JAPANESE PSYCHOLOGICAL ASSOCIATION

Japanese Psychological Research
2012, Volume 54, No. 1, 16–26
Special issue: Stereoscopic depth perception

doi: 10.1111/j.1468-5884.2011.00503.x

Isolation of two binocular mechanisms for motion in depth: A model and psychophysics[1]

SATOSHI SHIOIRI*, KAZUMICHI MATSUMIYA, and KAZUYA MATSUBARA
Tohoku University

Abstract: Random dot patterns (RDPs) have been used to stimulate in isolation either of two binocular mechanisms for motion in depth: interocular velocity difference (IOVD), and changing disparity over time (CDOT). First, we examined how these stimuli isolate either mechanism using models based on motion/disparity energy detection. In the model, the IOVD mechanism calculates the difference in motion signal between the two eyes, and the CDOT mechanism calculates the difference in disparity signal between sequential stereo images. The simulation revealed that uncorrelated (0% correlation) RDPs are useful to isolate either of the two mechanisms, while the contrast-reversed version (anticorrelation) of RDPs may not. Second, we compared the IOVD model predictions with experimental data from a previous study for motion in depth with various contrasts, displacements and vertical shifts between the two eyes. The simulation showed that the model predicts the general trends of the effects of contrast, displacement and vertical shift shown in the data. This suggests the physiological plausibility of the energy-based model of motion in depth.

Key words: motion in depth, interocular velocity difference, motion energy model, disparity energy model.

A variety of depth cues have been identified in retinal images. The cues include linear perspective, aerial perspective, texture gradient, binocular disparity, and others (Howard & Rogers, 1995). Despite the importance of 3-D motion in everyday life (e.g., for estimating time to contact; Lee, 1976; Matsumiya & Ando, 2009; Regan & Gray, 2000), there are few studies of motion in depth. Several cues can be used to see motion in depth (Shioiri, Morinaga, & Yaguchi, 2002), two of which are binocular cues (Brooks & Mather, 2000; Cumming & Parker, 1994; Harris & Sumnall, 2000; Regan, 1993; Shioiri, Saisho, & Yaguchi, 2000). We focus on binocular cues in this study.

One of the two binocular cues for perceiving motion in depth is the interocular velocity difference (IOVD) and the other is changing disparity over time (CDOT; or disparity change in time, DCT). Disparity is the difference between the left and right retinal images according to the distant differences of objects. The interocular velocity difference is the direction and/or speed difference between the motion of the left

*Correspondence concerning this article should be sent to: Satoshi Shioiri, Research Institute of Electrical Communication, Tohoku University, Katahira, Aoba-ku, Sendai 980-8577, Japan. (E-mail: shioiri@riec.tohoku.ac.jp)

[1]This study is partially supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT-JP), KAKENHI (Grants-in-Aid for Scientific Research (B)) 22330198 (2010) to S. Shioiri.

and right retinal images caused by the motion-in-depth of objects. Psychophysical studies have revealed that the visual system uses both of the cues. The use of the CDOT cue is supported by the finding that dynamic random dot stereograms, in which the disparity mechanism is assumed to be stimulated in isolation, provide perception of motion in depth (Cumming & Parker, 1994; Harris, Nefs, & Grafton, 2008; Julesz, 1971; Read, Parker, & Cumming, 2002). There was little support for the use of the IOVD cue (Cumming & Parker, 1994) before the end of the last century (Brooks & Mather, 2000) (Maeda, Sato, Ohmura, Miyazaki, Wang, & Awaya, 1999; Shioiri et al., 2000). For instance, Shioiri et al. (2000) showed that moving random dots in opposite directions between the two eyes provided motion-in-depth perception even when uncorrelated random dots were presented to the left and right retinas. Recent studies have confirmed this fact using different experimental paradigms (Brooks, 2001, 2002a,b; Fernandez & Farell, 2005, 2006; Rokers, Cormack, & Huk, 2008, 2009; Shioiri, Kakehi, Tashiro, & Yaguchi, 2009; Shioiri, Kakehi, & Yaguchi, 2002; Shioiri, Nakajima, Kakehi, & Yaguchi, 2008; Watanabe, Kezuka, Harasawa, Usui, Yaguchi, & Shioiri, 2008; see also Nefs & Harris, 2010).

Most of the researchers now agree that the two cues contribute to perception of motion in depth. However, it is an open question as to whether the results of an experiment indeed indicate the use of the IOVD or the CDOT cue for perceiving motion in depth. It is particularly important to examine whether an experimental condition stimulates one of the two mechanisms without stimulating the other. No stimulus has been confirmed to have only one of the two binocular cues. In this study, we used models of motion in depth to examine whether there is only one type of cue available in a stimulus or if there is an influence of the other cue.

There are two major techniques by which to isolate each of the two cues. First, Shioiri et al. (2000) controlled the dot correlation between the left and right eye images (binocular or spatial correlation) and the correlation between the frames in a two-frame apparent motion stimulus (temporal correlation) with random dot patterns (RDPs). To remove the CDOT cue, they used RDPs with 100% temporal and 0% binocular correlations, and to remove the IOVD cue they used RDPs with 0% temporal and 100% binocular correlations. These manipulations have a possible problem. The 0% correlation does not mean that there is no information available. This has been pointed out often as a problem for removing the CDOT cue (Harris et al., 2008), but the same criticism should apply for removing the IOVD cue (Shioiri et al., 2008). Careful experimental design is necessary for the use of these stimuli (Shioiri et al., 2000).

The second method is similar to the first one. It uses contrast-reversed versions of RDPs (anticorrelated or –100% correlated RDPs) instead of uncorrelated RDPs. The CDOT cues are removed if we assume that no useful disparity information remains in binocularly anticorrelated RDPs (Rokers et al., 2008; Rokers et al., 2009). Similarly, the IOVD cues are removed if we assume that no useful disparity information remains in temporally anticorrelated RDPs (no study has used this type of stimulus). While no serious analysis has been reported for the anticorrelated RDPs, studies of contrast reversals both for motion and depth perception suggest possible problems in cue removals. Contrast reversal could reverse the perceived direction of motion (Anderson & Burr, 1985; Anstis, 1970; Anstis & Rogers, 1975; Boulton & Baker, 1993; Braddick, 1980; Maruya, Mugishima, & Sato, 2003; Sato, 1989; Shioiri & Cavanagh, 1990; Takeuchi & De Valois, 1997, 2009) and that of depth (Anstis & Rogers, 1975; Ohzawa, 1998; Rogers & Anstis, 1975; Tanabe, Yasuoka, & Fujita, 2008).

The purpose of the present study is to investigate how the uncorrelated RDP and the contrast reversal (anticorrelated) RDP remove either of the two cues. We first show the model response to these stimuli, and second attempted to predict the experimental results of Shioiri et al. (2000) using the model.

Model

We used the IOVD and CDOT models proposed in a previous study (Shioiri et al., 2009) with a modification of contrast nonlinearity. The models were based on the detection of motion energy (Adelson & Bergen, 1985; Watson & Ahumada, 1985) and on that of disparity energy (Ohzawa, DeAngelis, & Freeman, 1990). The model consists of the motion or disparity energy detectors as subunits and combines their outputs to extract motion-in-depth signals in IOVD or CDOT analysis. In the IOVD model, monocular motion energy is calculated using a pair of spatial filters in each eye (low-level motion detector). The two filters have different phase properties and their outputs are summed after squaring (motion energy (ME)). The difference between the motion energy in two opposite directions is calculated as the velocity signal along the horizontal axis for each eye.

$$V_{eye} = LeftME - RightME, \tag{1}$$

where *LeftME* indicates leftward motion energy and *RightME* indicates rightward motion energy, V_{eye} indicates the velocity of either in the left or the right eye image. The signal of motion in depth, *IOVD*, is defined as the difference between the velocity signals for the two eyes.

$$IOVD = V_{left} - V_{right}, \tag{2}$$

where positive values correspond to approaching motion.

The model for CDOT detection is similar. Static disparity energy is calculated using a pair of spatial filters in each frame (disparity detector). The two filters have different phase properties and their outputs are summed after squaring (disparity energy (DE)). The difference between the disparity energy in two frames is calculated as the disparity signal at a moment.

$$D_t = NearDE - FarDE, \tag{3}$$

where *NearME* indicates near (crossed) disparity energy and *FarDE* indicates far (uncrossed) disparity energy. D_t indicates depth at a moment. The signal of motion in depth, *CDOT*, is defined as the difference between the depth signals between times (t1 and t2).

$$CDOT = D_{t2} - D_{t1}, \tag{4}$$

where positive values correspond to approaching motion as well.

To simulate the effect of contrast, we added a nonlinear response, which is known to be in the early vision, before inputting stimulus signals to the models. Because the outputs of the present models are proportional to the input contrast, the output function simply follows the contrast nonlinearity adopted. We used a Naka-Rushton-type equation for the contrast response function.

$$f(C) = R_{max} \cdot C^n / (C^n + C_{50}{}^n), \tag{5}$$

where C represents the input contrast, R_{max} represents the maximum obtainable response, n the exponent governing the steepness of the function, and C_{50} the half-saturation contrast, yielding a response of R_{max} /2. We used C_{50} of 0.1 and n of 1.5 following the nonlinearity estimated for the magnocellular LGN cells (Sclar, Maunsell, & Lennie, 1990) and the nonlinearity estimated psychophysically for the speed threshold (Shioiri, Ito, Sakurai, & Yaguchi, 2002), which is also similar to the contrast nonlinearity of human visual cortices as estimated by an fMRI study (Gardner, Sun, Waggoner, Ueno, Tanaka, & Cheng, 2005).

The choice of the spatial frequency tuning of the motion detector is not critical for the broadband stimuli such as random dot patterns. However, the choice is directly related to the ability to detect a displacement in the stimulus. The model detector is most sensitive to the displacement that is the same size as a quarter cycle of the wavelength of the grating with the peak spatial frequency. In the present simulation, the spatial characteristics of the motion and disparity detectors were fixed to be sensitive to displacements of approximately 4 min

(peak sensitivity at 1.75 c/deg). The spatial extent of the Gabor filters was chosen rather arbitrarily. The ratio (σ/f) between the spatial frequency (f) and the space constant (σ) was 0.3 unless mentioned otherwise. A ratio of 0.3 or a similar value was often used for the receptive field profile of cortical cells in the early vision (Ohzawa et al., 1990; Qian, Andersen, & Adelson, 1994).

All stimuli used here consisted of two regions with opposite motion signals: the upper half approaches and the lower half recedes (the direction was fixed in the simulation). The output strength of the motion in-depth process was estimated from the difference between the responses in the two regions.

Stimulus patterns

To predict the response of each of the motion-in-depth detectors, we calculated model outputs to several types of stimulation with RDP stimuli. All stimuli were a pair of RDPs: the left and right RDPs for each of two temporal frames. With such stimuli, apparent motion was seen due to a spatial displacement added between the frames. There were six conditions: V+D+, V+D0, V0D+, V+D–, V–D+ and V–D–, where V stands for the velocity cue and D stands for the disparity cue. A plus symbol indicates a positive correlation of dots between the two RDPs, a zero indicates no correlation and a minus symbol indicates a negative correlation or contrast reversal. The V+D+ stimulus contained both IOVD and CDOT cues. A random dot stereogram with a disparity was replaced by the same random dot stereogram but with a different disparity to mimic a movement in depth (from back to front or from front to back). The left and right images were exchanged in the present study to have the same disparity in opposite directions across the replacement. In the V+D0 stimulus, uncorrelated versions of RDPs were used between the two eyes to minimize the effect of the IOVD cue, while temporal correlation was kept as 100%. In the V–D+ stimulus, the contrast of the second frame was reversed so that the temporal correlation was perfectly negative or –100%, while the binocular correlation was kept as 100%. The V+D0 and V+D– stimuli were similar. The binocular correlation was 0 in the V+D0 stimulus and negative (–100%) in the V+D– stimulus. The V–D– stimulus was the stimulus with contrast reversal both binocularly and temporally. In all six conditions, the stimulus was designed to have dots approaching in the upper half and dots receding in the lower half in a square field. The field size was set to be $4.3^\circ \times 4.3^\circ$ or 128×128 pixels of $2' \times 2'$ to equate the condition of a psychophysical experiment.

For model predictions, the spatial properties of the motion/disparity detectors have to be chosen so that they can detect the spatial displacement added between the two images (binocular or temporal displacement) in simulation. The disparity detector is most sensitive to the binocular difference of a quarter cycle of the wavelength of the peak spatial frequency of the spatial filter, and the model motion detector is most sensitive to the temporal displacement of the same size. We used a peak spatial frequency of 1.75 c/deg so that the model detects the displacements used in the experiment of Shioiri et al. (2000), which were 4, 8, and 16 min of displacements. The detectors are most sensitive to the displacement of approximately 8 min.

Figure 1 shows the simulation results of the IOVD and CDOT models for each stimulus condition averaged over 10 times of calculation with different RDPs. The displacement size was fixed to 4 min here but different displacements provided similar results if the size was smaller than 16 min (not shown). The lighter parts in the figure indicate the approaching motion and the darker parts the receding motion, with mid gray as no motion. Figure 2 shows the relative responses of the model outputs for each stimulus. The model response was defined as the spatial average of the difference between the upper and lower fields. The average field, however, was limited to the central field (64×32 pixels for each half) so that potential errors due to wraparound at the border in the calculation did not influence (although the results are virtually identical for the full field data). The response is normalized so that it becomes unity for the IOVD response to the V+D+ stimulus.

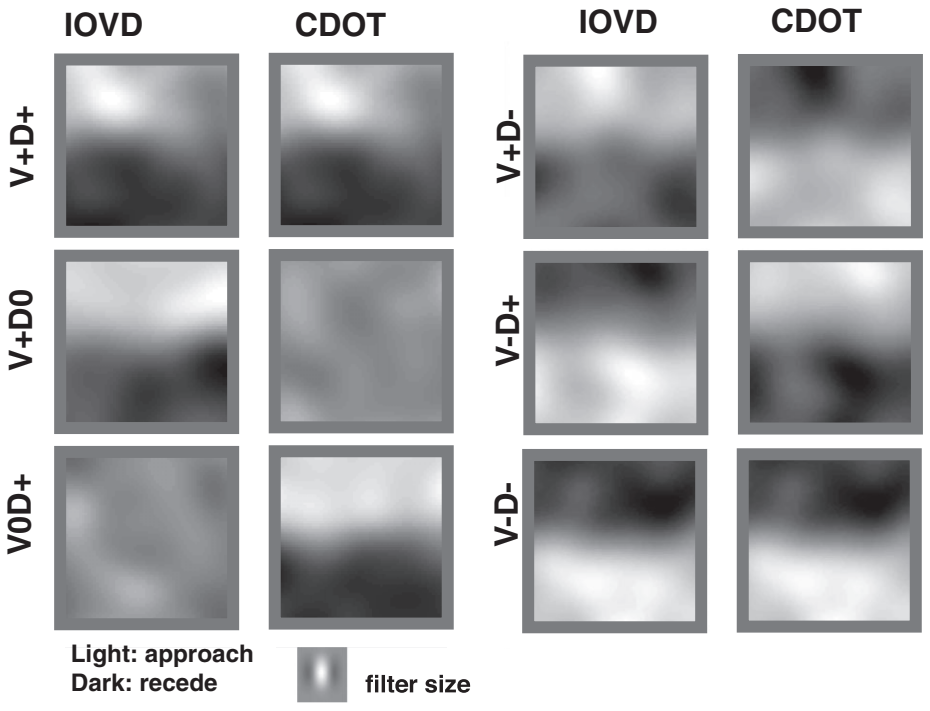

Figure 1 Simulation of the outputs of the interocular velocity difference (IOVD) and changing disparity over time (CDOT) models to the six types of random dot pattern (RDP) stimuli.

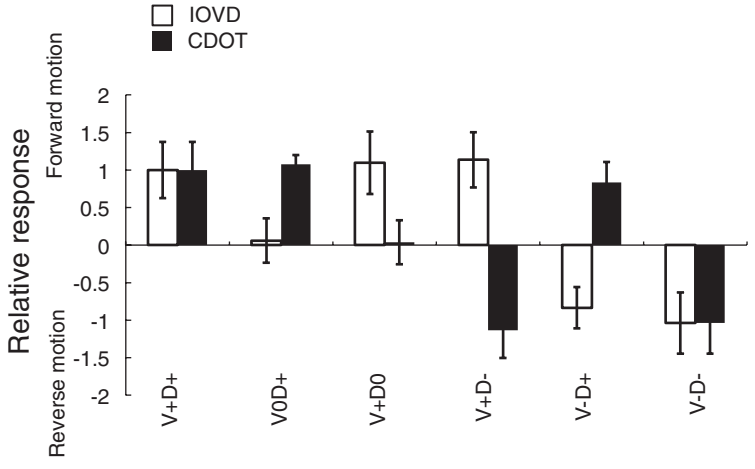

Figure 2 Relative responses of model simulations. The difference between the upper and lower halves was calculated for each simulated output to obtain an index for prediction. Positive values indicate motion in the simulated direction and negative values indicate reversed motion. Error bars represent standard error of mean over 10 random dot patterns. CDOT = changing disparity over time; IOVD = interocular velocity difference.

Both the IOVD and CDOT models detect the motion signal correctly in the V+D+ stimulus. Both the IOVD and CDOT models show a clear difference between the upper and lower halves of the stimulus field. Motion-in-depth signals can be detected by either mechanism. For the uncorrelated random dots, the IOVD model detects motion in the V+D0 stimulus and the CDOT model detects motion in the V0D+ stimulus. In contrast, the IOVD model cannot detect motion in the V0D+ stimulus and the CDOT model cannot detect motion in the V+D0 stimulus. These results suggest that two motion mechanisms can be isolated with the uncorrelated RDPs.

The results for the reverse contrast RDPs are different. The IOVD model detects motion in the V+D− stimulus and the CDOT model detects motion in the V−D+ stimulus, similarly to the V+D0 and V0D+ stimuli. However, the simulations showed that the IOVD model outputs motion signals in the opposite direction in the V−D+ stimulus and the V−D− stimulus, while the CDOT model outputs opposite motion signals in the V+D− stimulus and the V−D− stimulus. The reversal in the direction of motion in depth is consistent with the perceptual reversal of motion direction (Anderson & Burr, 1985; Anstis, 1970; Anstis & Rogers, 1975; Boulton & Baker, 1993; Braddick, 1980; Maruya et al., 2003; Sato, 1989; Shioiri & Cavanagh, 1990; Takeuchi & De Valois, 1997, 2009) and of depth direction (Anstis & Rogers, 1975; Ohzawa, 1998; Rogers & Anstis, 1975; Tanabe et al., 2008) with contrast reversal in stimulus patterns.

It should be noted that motion reversal is a consequence of narrow-band filtering in the models. Perceptual reversal is not surprising for periodic stimuli. Contrast reversal is equivalent to a half cycle displacement, and a quarter cycle displacement with contrast reversal is equivalent to a quarter cycle displacement in the opposite direction. However, the perceptual reversal in random dot is counterintuitive and a model is useful to understand the underlying mechanism (Shioiri & Cavanagh, 1990). The simulation predicts the perceptual reversal with a narrow-band spatial frequency filter. With a narrow-band frequency tuning, an RDP acts as if it were a pseudo-periodic pattern to the detectors in energy models. In other words, the motion reversal comes from using a Gabor function for spatial filters at the first stage of motion-in-depth analysis.

According to the present simulations, it is the uncorrelated stimulus that stimulates the IOVD or the CDOT mechanism in isolation. Measurements with the uncorrelated stimulus are expected to reflect the property of the mechanism isolated. The contrast-reversed stimulus, in contrast, can be used to confirm the existence of either of the motion mechanisms. Because contrast reversal realizes conditions where two mechanisms respond in opposite directions, the perceived direction indicates which of the two mechanisms is responsible for the perception. Perceiving motion in the V+D− condition, for example, cannot be attributed to any possible influence of disparity cues, as some researchers suggested (Cumming & Parker, 1994; Harris et al., 2008;) if the perceived direction is in the direction of the IOVD signal, as shown by Rokers et al. (2008, 2009).

Psychophysical experiments

We used the models for predicting the results of the psychophysical experiments of Shioiri et al. (2000). Their Experiment 1 measured performance varying the stimulus contrast and displacement size, and Experiment 2 varied the vertical offset between the left and right images. We calculated model responses while varying these factors. We describe their experiments shortly before showing the simulation results.

Experiment 1: Contrast and displacement

The stimuli were random dot patterns (50% light and 50% dark dots) that were binocularly uncorrelated, the same as for the V+D0 stimulus above. The stimuli contained two frames for each eye. 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. The upper half of the pattern moved leftward and the

lower half moved rightward or vice versa. The observer indicated the direction of motion in depth, and the percentages of correct responses were recorded for different contrast levels. Four observers participated in the experiment. The displacement size was either by 4, 8, or 16 min. The presentation duration of each frame was 480 ms. Each random dot field consisted of 256×256 pixels, which corresponds to 4.3 deg $\times$ 4.3 deg in the visual angle. Less pixel resolution was used in the model simulation (128×128 pixels) to save calculation time after checking that the difference did not cause any systematic differences.

Figure 3a shows the percentages of correct responses averaged over four observers as a function of stimulus contrast. Different symbols represent different displacement sizes. The performance is clearly higher than chance level for contrasts of 0.2 or higher for all displacement sizes and the percentages are similar for the 4 and 8 min displacement, while that for 16 min is lower.

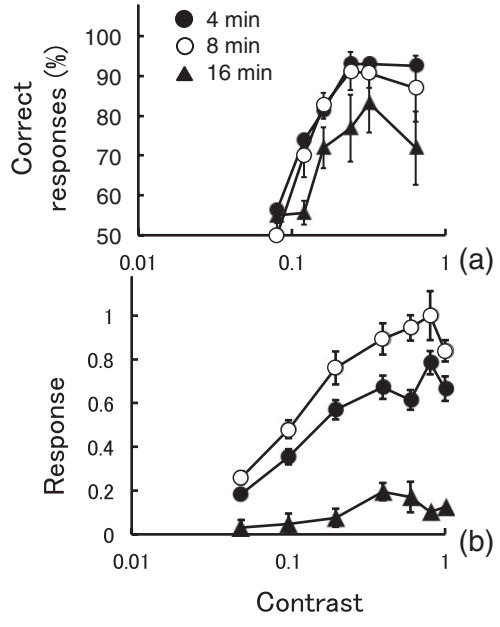

Figure 3 Percentage of correct responses in (a) a psychophysical experiment and (b) responses of model simulation for contrast changes. Error bars represent standard error of mean over (a) four observers or (b) 10 random dot patterns.

Experiment 2: Vertical offset

To avoid the random pairing of dots in binocularly uncorrelated RDPs, this experiment replaced parts of the random dot display with uniform gray horizontal bands that alternated in vertical position in the left and right images. As the bands with dots occupied alternate positions in the left and right images, there was no binocular overlap with appropriate binocular fusion (no-overlap condition). The band size varied between 8 and 20 min and the contrast of the dots was fixed to 0.32. In addition to this modified V+D0 stimulus, they used a V0D+ stimulus with similar modification to stimulate the disparity detection mechanism. Because of the gray band to remove the direct overlap of the dots, a vertical disparity with the same size as the band was introduced. Fixed displacement of 4 min was used in both conditions. Here, we also show the results of the overlap (or no-offset) condition, where there was no vertical shift while there were bands of dots (unpublished data corrected with the data reported in Shioiri et al.'s Experiment 2).

Figure 4a shows the percentages of correct responses as a function of band size in the no-overlap condition and Figure 4c shows the results in the overlap conditions. They are the results of three observers who participated in both the overlap and no-overlap conditions, instead of four reported in Shioiri et al. (one of the observer did not run the overlap condition). The results show that the direction of motion in depth is identified even without direct binocular overlaps of dots. The results of the overlap condition indicate that the band size itself does not change the performance in both the V+D0 and the V0D+ stimulus conditions.

Simulation

Contrast and displacement

Figure 3b shows the relative output of the IOVD model as a function of stimulus contrast for three displacement sizes. The value of one

corresponds to the maximum response among the responses in all conditions. The model captures the general trends of the experiment results. The similar effect of displacement between the model and the experiment suggests that the IOVD mechanism in the visual system is organized with spatial filters similar to those in the model. A similar effect of contrast suggests that the contrast nonlinearity of the motion system is also responsible for the nonlinearity of the motion in depth.

However, we should consider the effect of displacement size more in detail. We used the peak spatial frequency of 1.75 c/deg so that the model detects all the displacements of 4, 8, and 16 min. The model output for the 16 min displacement would be negative if the peak spatial frequency was higher than 1.9 c/deg (displacement has to be smaller than half a cycle of the sinusoidal component of the Gabor filter). Because of this restriction, the model cannot predict higher sensitivity to the 4 min displacement than to 8 min. The model with the peak spatial frequency of 1.75 c/deg senses 8 min displacement best because it is the closest to the displacement of a quarter of the cycle of the spatial frequency (8.6 min). If Gabor filters have a peak spatial frequency of 2.5 c/deg, the simulation shows a larger response to 4 min than to 8 min, as in the experiment results. However, the simulation to 16 min shows negative responses, indicating that the detector cannot detect the motion-in-depth signal in the condition. We expect that

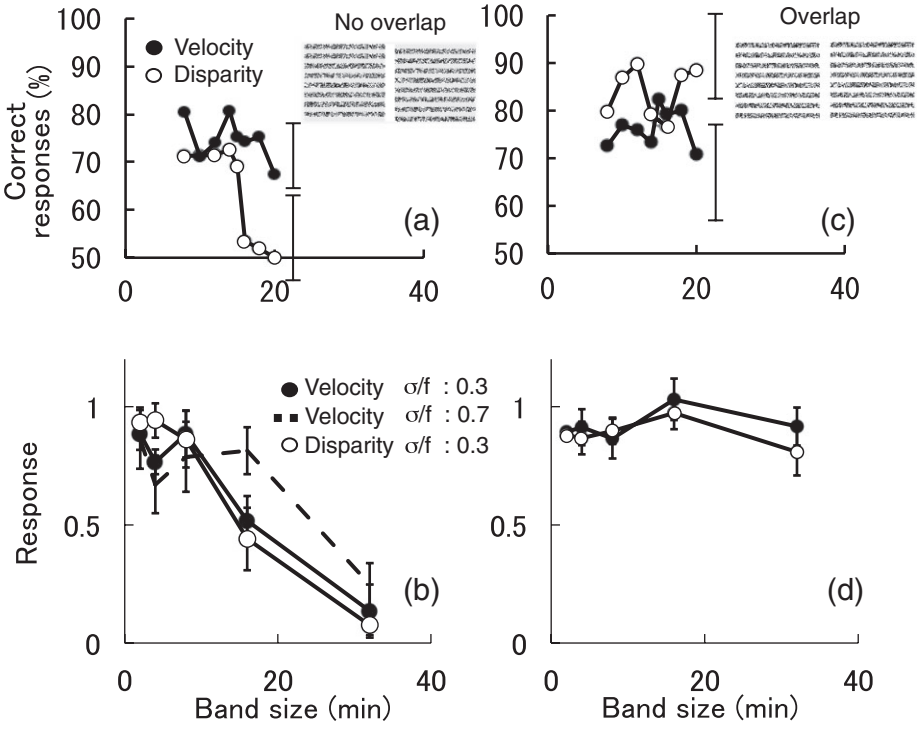

Figure 4 Percentage of correct responses in a psychophysical experiment: (a) for different vertical shifts (no-overlap condition) and (c) for different band sizes with no vertical shift (overlap condition). Responses of model simulation: (b) for different vertical shifts and (d) for different band sizes with no vertical shift. Error bars in (a) and (c) represent standard error of mean over three observers and error bars in (b) and (d) represent standard error of mean over 10 random dot patterns. The standard error averaged over all band sizes for each condition is plotted at the right of the rightmost data for the psychophysical results. Only representative error bars are given for clarity of the figures.

this problem can be solved in future study with multiple motion/disparity channels, as has been suggested for motion in depth (Shioiri et al., 2009).

Vertical offset

Figure 4b shows the results in the no-overlap condition and Figure 4d shows the results in the overlap condition. The model output for the no-overlap stimulus decreases with band size as in the experiment results, whereas no such decrease is shown for the overlap stimuli, also as the experiment results show. The model predictions are consistent with the experiment results in these aspects. There is also a difference between the model prediction and the experiment data. The experiment results show that performance declined more quickly with band size in the V0D+ stimulus condition than in the V+D0 stimulus condition. Having such results, Shioiri et al. speculated that the IOVD mechanism has a larger spatial integration area than the CDOT mechanism. If, for example, the CDOT signals are integrated within a 15 min area and if the IOVD signals are integrated within a 20 min or larger area, the effect of band size should be more severe in the V0D+ stimulus condition than in the V+D0 stimulus condition for the range of vertical shift in the experiment.

In the model, spatial integration size depends on the size of the spatial filters. As the same filter size was used for the IOVD and CDOT models, there is no particular reason to obtain differences in spatial integration between the two stimulus conditions. In order to examine whether the size difference in integration area explains the different effect of band size found between the two stimulus conditions, we executed simulations with larger space constants, σ, of Gabor filters for the IOVD model. The simulation showed a tendency that responses for larger bands increase with filters with larger space constants. The dashed line in Figure 4b shows the simulation result of the velocity condition with a σ/f of 0.7 instead of 0.3. The model predictions are similar to the experiment results (black circles in Figure 4a), showing relatively constant performance with band sizes of 20 min or smaller. Because the model predictions with σ/f of 0.3 are similar to the experiment results in the V0D+ condition (both showing quick decrease of response/performance at approximately 15 min), we suggest that the vertical extent of the receptive field of the IOVD mechanism is twofold or more larger than that of the CDOT mechanism.

Discussion

In order to investigate how different types of stimuli for motion in depth isolate either of the IOVD or the CDOT mechanisms, we built models of the IOVD and the CDOT mechanisms using motion and disparity energy detectors as subunits. Although the models do not predict the actual results expressed as a percentage of correct responses, they provide qualitative effects of stimulus conditions. The IOVD and CDOT models confirmed that removing the correlation from RDPs realizes stimulation of either mechanism in isolation. The stimulus without binocular correlations isolates the IOVD mechanism and the stimulus without temporal correlations isolates the CDOT mechanism. The models also showed that reversing contrast reverses motion direction in depth. Contrast reversal in one of the stereogram pair as well as contrast reversal of one of the two-frame motion pair reverses the direction of motion in depth. The experiment results with anticorrelated random dot patterns should be interpreted with consideration of the effect of motion energy in the reverse direction.

Our model does not include all of the possible factors which may contribute to perception of motion in depth. Particularly for the effect of contrast reversal, processes later than local energy detection probably play important roles. Depth perception in contrast-reversed stereogram (or anticorrelated stereogram) appears to be dependent on stimulus conditions (Cumming & Parker, 1997; Masson, Busettini, & Miles, 1997; Tanabe et al., 2008). Depth reversal in random dot stereograms can be seen with a reference plane to provide relative disparity, whereas only unstable cloud of dots may be

perceived without such a plane (Tanabe et al., 2008). The effect of the reference plane is likely due to processes later than disparity detection. If perception of motion in depth is suppressed at a higher stage in reversed contrast conditions (that is, if no motion in depth is seen), contrast-reversed RDPs may also isolate either of the IOVD or the CDOT mechanisms as do uncorrelated RDPs.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, **2**, 284–299.

Anderson, S. J., & Burr, D. C. (1985). Spatial and temporal selectivity of the human motion detection system. *Vision Research*, **25**, 1147–1154.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, **10**, 1411–1430.

Anstis, S. M., & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision Research*, **15**, 957–961.

Boulton, J. C., & Baker, C. L., Jr (1993). Dependence on stimulus onset asynchrony in apparent motion: Evidence for two mechanisms. *Vision Research*, **33**, 2013–2019.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **290**, 137–151.

Brooks, K. (2001). Stereomotion speed perception is contrast dependent. *Perception*, **30**, 725–731.

Brooks, K. R. (2002a). Interocular velocity difference contributes to stereomotion speed perception. *Journal of Vision*, **2** (3), 218–231.

Brooks, K. R. (2002b). Monocular motion adaptation affects the perceived trajectory of stereomotion. *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 1470–1482.

Brooks, K., & Mather, G. (2000). Perceived speed of motion in depth is reduced in the periphery. *Vision Research*, **40**, 3507–3516.

Cumming, B. G., & Parker, A. J. (1994). Binocular mechanisms for detecting motion-in-depth. *Vision Research*, **34**, 483–495.

Cumming, B. G., & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, **389**, 280–283.

Fernandez, J. M., & Farell, B. (2005). Seeing motion in depth using inter-ocular velocity differences. *Vision Research*, **45**, 2786–2798.

Fernandez, J. M., & Farell, B. (2006). Motion in depth from interocular velocity differences revealed by differential motion aftereffect. *Vision Research*, **46**, 1307–1317.

Gardner, J. L., Sun, P., Waggoner, R. A., Ueno, K., Tanaka, K., & Cheng, K. (2005). Contrast adaptation and representation in human early visual cortex. *Neuron*, **47**, 607–620.

Harris, J. M., Nefs, H. T., & Grafton, C. E. (2008). Binocular vision and motion-in-depth. *Spatial Vision*, **21**, 531–547.

Harris, J. M., & Sumnall, J. H. (2000). Detecting binocular 3D motion in static 3D noise: No effect of viewing distance. *Spatial Vision*, **14**, 11–19.

Howard, I. P., & Rogers, B. (1995). *Binocular vision and stereopsis*. Oxford: Oxford University Press.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago, IL: University of Chicago Press.

Lee, D. N. (1976). Theory of visual control of braking based on information about time-to-collision. *Perception*, **5**, 437–459.

Maeda, M., Sato, M., Ohmura, T., Miyazaki, Y., Wang, A. H., & Awaya, S. (1999). Binocular depth-from-motion in infantile and late-onset esotropia patients with poor stereopsis. *Investigative Ophthalmology and Visual Science*, **40**, 3031–3036.

Maruya, K., Mugishima, Y., & Sato, T. (2003). Reversed-phi perception with motion-defined motion stimuli. *Vision Research*, **43**, 2517–2526.

Masson, G. S., Busettini, C., & Miles, F. A. (1997). Vergence eye movements in response to binocular disparity without depth perception. *Nature*, **389**, 283–286.

Matsumiya, K., & Ando, H. (2009). World-centered perception of 3D object motion during visually guided self-motion. *Journal of Vision*, **9** (15), 11–13.

Nefs, H. T., & Harris, J. M. (2010). What visual information is used for stereoscopic depth displacement discrimination? *Perception*, **39**, 727–744.

Ohzawa, I. (1998). Mechanisms of stereoscopic vision: The disparity energy model. *Current Opinion in Neurobiology*, **8**, 509–515.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, **249**, 1037–1041.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994). Transparent motion perception as detection of unbalanced motion signals. III. Modeling. *Journal of Neuroscience*, **14**, 7381–7392.

Read, J. C., Parker, A. J., & Cumming, B. G. (2002). A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images. *Visual Neuroscience*, **19**, 735–753.

Regan, D. (1993). Binocular correlates of the direction of motion in depth. *Vision Research*, **33**, 2359–2360.

Regan, D., & Gray, R. (2000). Visually guided collision avoidance and collision achievement. *Trends in Cognitive Sciences*, **4**, 99–107.

Rogers, B. J., & Anstis, S. M. (1975). Reversed depth from positive and negative stereograms. *Perception*, **4**, 193–201.

Rokers, B., Cormack, L. K., & Huk, A. C. (2008). Strong percepts of motion through depth without strong percepts of position in depth. *Journal of Vision*, **8** (6), 1–10.

Rokers, B., Cormack, L. K., & Huk, A. C. (2009). Disparity- and velocity-based signals for three-dimensional motion perception in human MT+. *Nature Neuroscience*, **12**, 1050–1055.

Sato, T. (1989). Reversed apparent motion with random dot patterns. *Vision Research*, **29**, 1749–1758.

Sclar, G., Maunsell, J. H., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research*, **30**, 1–10.

Shioiri, S., & Cavanagh, P. (1990). ISI produces reverse apparent motion. *Vision Research*, **30**, 757–768.

Shioiri, S., Ito, S., Sakurai, K., & Yaguchi, H. (2002). Detection of relative and uniform motion. *Journal of the Optical Society of America: Optics, Image Science, and Vision*, **19**, 2169–2179.

Shioiri, S., Kakehi, D., Tashiro, T., & Yaguchi, H. (2009). Integration of monocular motion signals and the analysis of interocular velocity differences for the perception of motion-in-depth. *Journal of Vision*, **9** (10), 11–17.

Shioiri, S., Kakehi, D., & Yaguchi, H. (2002). Motion in depth perception based on monocular motion aftereffect. Second Asian Conference on Vision. Gyeongju, Korea, 10.

Shioiri, S., Morinaga, A., & Yaguchi, H. (2002). Depth perception of moving objects. In B. Javidi & F. Okano (Eds.), *3D television, video and display technology*. Berlin: Springer-Verlag, pp. 397–427.

Shioiri, S., Nakajima, T., Kakehi, D., & Yaguchi, H. (2008). Differences in temporal frequency tuning between the two binocular mechanisms for seeing motion in depth. *Journal of the Optical Society of America: Optics, Image Science, and Vision*, **25**, 1574–1585.

Shioiri, S., Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research*, **40**, 2565–2572.

Takeuchi, T., & De Valois, K. K. (1997). Motion-reversal reveals two motion mechanisms functioning in scotopic vision. *Vision Research*, **37**, 745–755.

Takeuchi, T., & De Valois, K. K. (2009). Visual motion mechanisms under low retinal illuminance revealed by motion reversal. *Vision Research*, **49**, 801–809.

Tanabe, S., Yasuoka, S., & Fujita, I. (2008). Disparity-energy signals in perceived stereoscopic depth. *Journal of Vision*, **8** (3) 22, 1–10.

Watanabe, Y., Kezuka, T., Harasawa, K., Usui, M., Yaguchi, H., & Shioiri, S. (2008). A new method for assessing motion-in-depth perception in strabismic patients. *British Journal of Ophthalmology*, **92**, 47–50.

Watson, A. B., & Ahumada, A. J., Jr (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, **2**, 322–341.

(Received May 12, 2011; accepted November 5, 2011)