The dynamics of the visual system in combining conflicting KDE and binocular stereopsis cues

KEN'YA UOMORI

ATR Auditory and Visual Perception Research Laboratories, Kyoto, Japan

and

SHIN'YA NISHIDA

NTT Basic Research Laboratories, Tokyo, Japan

The dynamics of the visual system in combining multiple depth cues were investigated by measuring the temporal change in the perceived 3-D shape of a random-dot stimulus with conflicting kinetic depth effect (KDE) and binocular stereopsis cues. The KDE shape perception dominated for the first few seconds, and then was gradually supplanted by the stereo shape perception. The effects of various pre-adaptation stimuli suggested that the temporal change in the perceived shape resulted from a self-adaptation of the KDE mechanism that occurs mainly at the levels of motion and relative motion detection.

There are many cues for depth perception, such as binocular stereopsis, object motion (kinetic depth effect, or KDE), observer's motion (motion parallax), shading, and texture gradient. Our visual system recovers the 3-D structures of the outside world by combining depth estimations from these cues. To investigate how various depth cues are combined, several researchers have used stimuli that involve conflicting depth cues at the same position (Bülthoff & Mallot, 1990; Landy, Maloney, & Young, 1991; Maloney & Landy, 1989; Rogers & Collett, 1989). For example, Maloney and his colleagues (Landy et al., 1991; Maloney & Landy, 1989) examined depth perception with conflicting KDE and texture gradient cues and postulated that the final depth perception from multiple cues is based on a linear summation of the depths estimated from each cue, unless those estimates are strongly inconsistent. The coefficient of each cue is based on the cue's reliability, as derived from its consistency with the other cues.

These studies implicitly assume stable depth perception for stimuli with conflicting depth cues. However, after continuous observation of a display showing conflicting KDE and stereo cues, we noticed that the depth perception changes over time, even though the stimuli do not change. At first, the depth perception determined by the KDE cue dominated, but then the depth perception determined by the stereo cue gradually appeared. We felt

that clarification of the temporal sequence of this depth change was important in obtaining a proper estimate of the contribution of each cue to the final depth perception. We also expected that the examination of this phenomenon might provide new insights into the dynamics of the stereopsis system, the KDE system, or a system combining them.

In the present paper we examine the temporal change in depth perception for a conflicting-cue stimulus. In the first experiment we continuously monitored the subjects' depth perception to analyze the temporal sequence of the phenomenon quantitatively. In the subsequent experiments we introduced various preadaptation stimuli in order to investigate the mechanism of the temporal change in depth perception. Our results suggest that the shape change results from a self-adaptation of the KDE mechanism that occurs mainly at the levels of motion and relative motion detection.

EXPERIMENT 1

Temporal Change in Depth Perception

Method

Stimuli and Apparatus. We made a conflicting-cue stimulus in the following way. First, we created a sequence of images that simulated a view of random dots lying on the surface of an opaque vertical cylinder rotating around the vertical axis (Figure 1a), assuming an orthographic projection. The image subtended 200° (vertical) $\times$ 201° (horizontal), with one dot subtending $1^\circ \times 1^\circ$. Each dot was displayed as a white pixel on a dark background. The image was updated every 50 msec (1 frame = 3 vertical synchronization intervals of the display), with no interstimulus interval (ISI). The angular velocity of the cylinder was 40 deg/sec. To eliminate the dot density gradient that might serve as a depth cue, we set the lifetime of each dot at 10 frames (500 msec) and controlled the probability of dot generation to keep the dot density at about 2.5% at all horizontal positions (Doshier, Landy, & Sperling, 1989; Treue, Husain, & Andersen, 1991). Desynchronized dot generation in com-

All of the experiments reported here were conducted at the ATR Auditory and Visual Perception Research Laboratories. The authors are grateful to the laboratory's president, Eiji Yodogawa, and the department head, Keiichi Ueno, for the opportunity to conduct this research. We also thank Takao Sato (NTT), Farley Norman (Ohio State University), and Michael Landy (New York University), who gave us helpful advice. Address correspondence to K. Uomori, Central Research Laboratories, Matsushita Electric Industrial Co. Ltd., 3-4 Hikari-dai, Seika-cho, Soraku-gun Kyoto 619-02, Japan.

Figure 1. Schematic illustration of the shapes suggested by the conflicting-cue stimulus. (a) Vertical cylinder (KDE cue). (b) Horizontal cylinder (binocular stereopsis cue).

Figure 2. Temporal sequence of trials in Experiments 1 and 2. (a) The conflicting-cue stimulus was presented continuously for 60 sec (K+S condition in Experiment 1). (b) Presentation of pre-adaptation stimulus with stereo cue was followed by presentation of the conflicting-cue stimulus (S→K+S condition in Experiment 2). (c) Presentation of pre-adaptation stimulus with KDE cue was followed by presentation of the conflicting-cue stimulus (K→K+S condition in Experiment 2).

bination with a short lifetime allowed us to make a continuously rotating cylinder with a repetition cycle of 10 frames.¹

Next, we added a conflicting stereo cue to the sequence of images. We presented two sequences of images side by side, with stereo disparities that simulated a horizontal cylinder (Figure 1b). The disparity ranged between $-4'$ and $4'$ around the fixation plane.

The stimulus was presented on the face of a CRT display (Ikegami C/T-20; vertical synchronization cycle = 60 Hz) controlled by a workstation (HP9000 S400 TurboVRX). The subject viewed the display binocularly through a haploscope with a chinrest, at a viewing distance of 91 cm. A fixation point was located at the center of each stimulus.

Procedure. We measured the temporal change in depth perception for two depth cues (binocular stereopsis and KDE) with conflicting 3-D shapes. In one trial the conflicting-cue stimulus was presented continuously for 60 sec (Figure 2a). The subject's task was to report the perceived 3-D shape by pressing one of three mouse buttons during the stimulus presentation. The choices were (1) the shape determined by binocular stereopsis (horizontal cylinder), (2) the shape determined by KDE (vertical cylinder), and (3) neither of these, or an intermediate shape. The sampling frequency of the subject's responses was 60 Hz. During stimulus presentation, the subject was told to fixate on a cross that was located at the center of the stimulus. To reduce the intertrial adaptation effect, the rotation direction was reversed at every new trial, with at least 3 min between trials. The data for Experiment 1 and the two conditions of Experiment 2 were collected simultaneously. These three conditions were randomly organized in a single block of three trials for each condition. Nine trials were completed in each condition.

Subjects. Four subjects participated in Experiment 1. Two of them (the authors K.U. and S.N., both males) were myopic, with their acuity corrected by glasses or contact lenses. The others (Subjects N.O. and K.M., both females) had normal vision and were naive regarding the purpose of this study. All the subjects had normal stereo vision.

Results

Figure 3 shows the results for Subject S.N. The probabilities of each choice (shape from binocular stereopsis, shape from KDE, or an intermediate shape), calculated from the data of nine trials, are plotted as functions of time. The data are smoothed temporally by averaging within a rectangular window of 3 sec. Point T in Figure 2 corresponds to 0 sec in Figure 3. The figure shows that the subject frequently reported the KDE shape for the first

Figure 3. Temporal change in probability for each response category obtained in Experiment 1 (K+S condition). The data (Subject S.N.) are smoothed temporally with a rectangular window of 3 sec.

20 sec. Thereafter, the intermediate-shape response increased. According to verbal reports, the subjects perceived a shape that would be given by a linear summation of vertical and horizontal cylinders (a barrel shape), rather than the two cylinders simultaneously. In the last 30 sec, Subject S.N. reported the stereo shape as frequently as the intermediate shape. These tendencies suggest that the relative contribution of the two depth cues changed over time.

To further clarify this, we calculated a response value at each time by assigning the values of -1 to the response for the binocular stereopsis shape, 1 to the response for the KDE shape, and 0 to the response for the intermediate shape, and averaging the values over nine trials. (Here we assumed an interval scale for the three categories of response.) A response value of 1.0 indicates perfect KDE shape perception (vertical cylinder) over the trials, and a value of -1.0 indicates perfect binocular stereopsis shape perception (horizontal cylinder). The thick lines in Figure 4 show the temporal change in response values

temporally smoothed with a 3-sec rectangular window. The results for 3 subjects (K.U., S.N., and K.M.) consistently show a temporal change in depth perception. At first, their response values are close to 1.0 , but they steadily decrease with time. This implies that these subjects tended to perceive shape on the basis of KDE for the first several seconds, but then their perception gradually changed to favor the binocular stereopsis shape. The results for Subject N.O. show this tendency only slightly.

Examined closely, the depth perception reported during the latter half of the trial differs slightly among subjects. The response of Subject K.U. converges to a value of about zero. The responses of S.N. and K.M. converge to values of around -0.3 and -0.5 , respectively. The response of Subject N.O. converges to a value of about 0.5 . This individual variance probably reflects a difference in the relative strength of the two depth cues and/or the difficulty in setting a stable response criterion across subjects. For Subject N.O., whose data show little temporal change, a strong contribution of the KDE cue and/or her

Figure 4. Temporal change in perceived shape for conflicting-cue stimulus under the three conditions shown in Figure 2. The response value was calculated by assigning -1 to the response for the binocular stereopsis shape, 1 to the response for the KDE shape, and 0 to the response for neither of these or an intermediate shape, and averaging them over all trials. The data are smoothed temporally with a rectangular window of 3 sec.

response criterion favoring the KDE shape might reduce the degree of temporal change.

Several studies have shown anisotropies in the perceived curvature of surfaces specified by both motion and binocular stereopsis. For stereo perception, curvature sensitivity in the vertical direction is better than that in the horizontal direction (Norman, Todd, Tittle, & Perotti, 1993; Rogers & Graham, 1983), and for KDE, sensitivity is better for curvature along a direction parallel to the axis of rotation than for curvature along a direction perpendicular to the axis of rotation (Cornilleau-Pérès & Droulez, 1989; Norman & Lappin, 1992; Norman et al., 1993). In Experiment 1, we used stereo curvature in the vertical direction and KDE curvature along a direction perpendicular to the axis of rotation, which implies that the stereo curvature was optimal, but the KDE curvature was suboptimal.

To test the effects of anisotropies on the temporal change in shape perception for the conflicting-cue stimulus, we conducted a control experiment with 1 subject (K.U.), using a stimulus that comprised a velocity field simulating a view of a horizontal cylinder rotating around the horizontal axis and binocular stereo disparity simulating a vertical cylinder. We called this stimulus *V-flow*, and the original one *H-flow*. The *V-flow* had suboptimal curvatures for both KDE and stereo. In the experiment, the data for the *H-flow* were collected again, and the two kinds of stimuli were randomly presented within a session. The results are shown in Figure 5. The response value for the *V-flow* shows temporal change that is similar to the *H-flow*: the KDE shape dominated at first, then the contribution of the stereo shape gradually increased. A notable difference between the two conditions is that the strong contribution of the KDE shape persisted longer for the *V-flow* than for the *H-flow*. This tendency was

Figure 5. Temporal change in perceived shape for two conflicting-cue stimuli. *H-flow* was the stimulus shown in Figure 1. *V-flow* was a stimulus consisting of a velocity field that simulated a horizontal cylinder rotating around the horizontal axis and binocular stereo disparities that simulated a vertical cylinder. The data (Subject K.U.) are smoothed temporally with a rectangular window of 3 sec.

confirmed by an informal observation with another subject. This difference is consistent with the anisotropy reported earlier, which predicts that the contribution of the stereo cue is less for the *V-flow*. Thus, the anisotropies of stereo and KDE perception may quantitatively affect the result, but it is unlikely that they change the qualitative nature of our finding.

Discussion

The results of Experiment 1 showed that the depth perception for a stimulus with conflicting KDE and stereo cues changed temporally: first the KDE cue dominated, then the effect of the stereo cue increased. In Figure 4, one can notice a slight fluctuation of the response value, which reflects oscillative changes in the shape percept. Examining the result of each trial, we found that the percept flipped—sometimes between the KDE and intermediate shapes, sometimes between the intermediate and stereo shapes, and sometimes even between the KDE and stereo shapes. Some of these flips might have been caused by involuntary blinking or eye movement. Even voluntary attention shift might have played some role. When the data were averaged over the trials, however, the oscillative change in shape became less visible, and the general trend of a shift from the KDE shape to the stereo shape became clear.

In one stimulus (*H-flow*), the stereo curvature was in the vertical direction, and the KDE curvature was along a direction perpendicular to the axis of rotation. Under this condition, Norman et al. (1993) reported that the stereo shape dominated the KDE shape, whereas we found, at least for the initial few seconds, that the KDE shape dominated the stereo shape. This difference probably just reflects our choice of stimulus parameters (e.g., simulated shape, velocity, and size of binocular disparity) that favor KDE perception. It should be noted that we did not attempt to address stimulus factors here, such as curvature orientation and simulated shape, which may affect the strength of each depth cue, but were concerned instead with the temporal change of shape perception for a conflicting-cue stimulus consisting of a given combination of stimulus parameters.

Why does the temporal change occur for our conflicting-cue stimulus? There are at least two possible hypotheses.

1. The effect of the stereo cue increases over time. Binocular stereopsis processing may require several seconds to match the left and right images. In addition, after completion of image matching, the output of a binocular stereopsis system may gradually become strong or stable.

2. The effect of the KDE cue decreases over time. It is widely known that prolonged presentation of a constant stimulus reduces the response of the visual system to that stimulus. Assuming that self-adaptation is much stronger or quicker to the KDE cue than to the stereo cue, the relative effect of the KDE cue will gradually decrease.

To test these hypotheses, we introduced two kinds of pre-adaptation stimuli in Experiment 2. One was a stim-

ulus with only a stereo cue, and the other was a stimulus with a KDE cue presented on the fixation plane. If the first hypothesis is valid, pre-adaptation to the stereo cue will enhance the contribution of the stereo cue to the depth perception for our conflicting-cue test stimulus. If the second hypothesis is valid, pre-adaptation to the KDE cue will reduce the contribution of the KDE cue.

EXPERIMENT 2

Pre-Adaptation to Stereo and KDE Stimuli

Method

The same methods were employed as in Experiment 1, except for the use of pre-adaptation stimuli. Under one condition ($S \rightarrow K+S$), we employed a pre-adaptation stimulus that had a disparity field identical to that of the test stimulus (horizontal cylinder) and consisted of uniformly moving (1 deg/sec) random dots. Under the other condition ($K \rightarrow K+S$), the pre-adaptation stimulus was a sequence of random dots simulating a rotating vertical cylinder, presented on the zero-disparity plane.

For each trial, one of the pre-adaptation stimuli was presented for 10 sec, followed by a 60-sec presentation of the conflicting-cue stimulus (the H-flow test stimulus). Under the $S \rightarrow K+S$ condition, the pre-adaptation stimulus was gradually replaced by the test stimulus during the last 500 msec of the pre-adaptation period in order to reduce the subject's impression of a sudden change in stim-

ulus (Figure 2b). Under the $K \rightarrow K+S$ condition, the pre-adaptation stimulus was replaced by the test stimulus, with no transition time (Figure 2c) because the disparity change was small. As noted above, the data for these two conditions were collected along with the data for the previous experiment. A fixation point was located at the center of each stimulus.

Results

The results for the $S \rightarrow K+S$ and $K \rightarrow K+S$ conditions (dotted and thin lines, respectively, in Figure 4) show a temporal change in perceived shape that is similar to that found in Experiment 1 ($K+S$ condition). First, the KDE shape dominates, then the effect of the stereo cue increases. This tendency is clearly observed, even for Subject N.O., who showed only a slight temporal change in Experiment 1.

The response value for the $S \rightarrow K+S$ condition is roughly the same as the value for the $K+S$ condition. Compared with these two conditions, the response value for the $K \rightarrow K+S$ condition is small, especially for the initial few seconds, which indicates a decrease in the contribution of the KDE cue. To clarify these tendencies, we averaged the response value over the first 20 sec for each condition (Figure 6). Analysis of variance (ANOVA) for the three adaptations and 4 subjects revealed significant

Figure 6. Response value for three conditions ($K+S$, $K \rightarrow K+S$, $S \rightarrow K+S$) averaged over the first 20 sec of the test period. Error bars show ± 1 standard error among trials. The vertical scale is adjusted to the response range of each subject.

main effects of pre-adaptation [$F(2,96) = 4.97, p < .01$] and subject [$F(3,96) = 10.23, p < .01$]. The interaction between pre-adaptation and subject was not significant ($F < 1.0$). Further probing of the mean differences for pre-adaptations (Tukey-Kramer's *HSD* test, using the JMP from SAS Institute Inc., a statistical analysis software on Macintosh), indicated significant differences between the K+S and K→K+S conditions and between the S→K+S and K→K+S conditions, but no significant difference between the K+S and S→K+S conditions (5% level). Thus, the average values were roughly the same for the K+S and S→K+S conditions, but the value for the K→K+S condition was significantly smaller than the other two.

Discussion

Under the S→K+S condition, it was expected that matching the left and right images by binocular stereopsis processing would be completed within the 10-sec presentation of the pre-adaptation stimulus, which contained a stereopsis cue that was identical to that of the conflicting-cue test stimulus. This expectation was confirmed by the subjects' verbal reports that the horizontal cylinder was perceived during the pre-adaptation period. If the temporal change in shape perception for our conflicting-cue stimulus is due to the rather long period of time required by the initial stereopsis processing, then the effect of the stereo cue is stronger at the start of the test period under the S→K+S condition than under the K+S condition. A similar tendency is predicted if we assume that the effect of the stereo cue gradually increases due to other, unknown factors. However, we did not find a stronger contribution of the stereo cue for the S→K+S condition.

The second hypothesis, suggested previously, was that the temporal decrease of the effect of the KDE cue is due to the quick self-adaptation of the KDE processing system. According to this hypothesis, the relative contribution of the KDE cue is reduced for the K→K+S condition because the self-adaptation of the KDE system is expected to proceed during the pre-adaptation period. The results obtained here are quite consistent with this prediction. The average response value for the K→K+S condition was smaller than for the other two (Figure 6). Furthermore, the results of some of the subjects showed that the response function for the K→K+S condition might be described as a replica of the other two functions by shifting them to the left about 10 sec (Figure 4). Therefore, it is highly likely that the temporal change in the perceived shape of our conflicting-cue stimulus is mainly due to the self-adaptation of the KDE processing system.

EXPERIMENT 3 Cross-Adaptation Experiment

The results of the previous experiments suggest that the temporal change in depth perception for the conflicting-cue stimulus mainly results from self-adaptation to the KDE cue. The next issue we addressed was the processing level at which that self-adaptation occurs. We tried

Figure 7. A hierarchical model of the KDE mechanism assumed in the examination of adaptation levels. See text for further explanation.

to find this answer by introducing various pre-adaptation stimuli. In the selection of the pre-adaptation stimuli, we assumed a three-stage model for the KDE processing (Figure 7). In the first stage, local motion is detected at every retinal position by a mechanism similar to that proposed by Van Santen and Sperling (1985) and Adelson and Bergen (1985). Next, local relative motion is calculated from the outputs of the local motion detectors. The existence of a relative motion detector has been suggested by several psychophysical and neurophysiological studies (Allman, Miezin, & McGuinness, 1985; Loomis & Nakayama, 1973; Regan, 1986). It is likely that the relative motion detector plays a significant role in the KDE processing, because the complex relationships among image velocities at various points contain information that is critical in recovering the object's structure from motion (Droulez & Cornilleau-Pérès, 1990). Finally, a 3-D structure is reconstructed by synthesizing the outputs of the local motion and local relative motion detectors. Several possible algorithms have been proposed for this stage (Ando, 1991; Koenderink & Van Doorn, 1991; Ullman, 1979), but specification of the algorithm is beyond the scope of this study.

Method

The methods employed in Experiment 3 were somewhat different from those of the previous experiments.

The conflicting-cue test stimulus was slightly changed. The dot size was $2' \times 2'$, and the dot density was 2.0%. The maximum

horizontal disparity was 10'. The lifetime of the dot remained at 10 frames, but the repetition cycle of the stimulus was increased to 20 frames.

Seven pre-adaptation stimuli were employed. Unless otherwise noted, the parameters of the pre-adaptation stimuli (e.g., stimulus size, moving direction, and position) were the same as for the test stimulus, and a fixation point was located at the center of the stimulus. All the pre-adaptation stimuli, which are listed below, were presented on the fixation (zero-disparity) plane.

A. Still, random-dot stimulus: still image (SI) condition (Figure 8a). We used this condition as a control.

B. Moving random-dot stimulus, in which all dots had a uniform velocity of 1 deg/sec: uniform velocity field (UVF) condition (Figure 8b).

C. Moving random-dot stimulus, in which the dot velocity was the same as in the KDE stimulus, but dot position was randomized within the stimulus field: random velocity field (RVF) condition (Figure 8c).

D. KDE stimulus that simulates the view of a rotating vertical cylinder: KDE condition (Figure 8d). This is identical to the test stimulus, except for the disparity setting.

E. KDE stimulus, whose position was shifted to the right (or left) by one half of the stimulus size: position change (KDE-P) condi-

Figure 8. Variety of pre-adaptation stimuli used in Experiment 3. See text for further explanation.

Figure 9. Velocity fields of pre-adaptation stimuli for the KDE condition [(d) in Figure 8], and for the DC-S condition [(g) in Figure 8].

Figure 10. Temporal sequence of a trial in Experiment 3.

tion (Figure 8e). In this case, a fixation point was located on a side edge of the stimulus.

F. KDE stimulus, whose rotation direction was opposite that of the test stimulus: direction change (KDE-D) condition (Figure 8f).

G. Moving random-dot stimulus, in which the velocity of each dot was the same as that of a dot at the same horizontal position in the KDE stimulus minus the maximum velocity of the KDE stimulus: DC-subtracted (DC-S) condition (Figure 8g). Figure 9 shows the velocity distributions of the KDE and DC-S conditions. The DC-S stimulus was sometimes perceived as a 3-D structure that was quite different from a rotating cylinder. Note that the dots in the DC-S stimulus moved in the direction opposite to those in the test stimulus, because we made the rotation direction of the original KDE stimulus the same as that of the test stimulus.

One trial consisted of a 20-sec presentation of one of the pre-adaptation stimuli, followed by a 20-sec presentation of the test stimulus (Figure 10). The stimuli were changed abruptly. As before, the subjects' task was to report the perceived shape by pressing a mouse button during the test period. Generally, one block consisted of a random presentation of seven trials, one for each of the pre-adaptation conditions. Ten trials were conducted for each condition.

The 2 authors and 2 naive subjects (K.M. and H.I.) participated in this experiment. Subject H.I. (male) was myopic, with his acuity corrected by glasses.

Results

As in the previous experiments, we analyzed the temporal change of a subject's responses by assigning three values (-1, 1, or 0) to the three response choices (i.e., binocular stereopsis shape, KDE shape, and intermediate shape). Typical results are shown in Figure 11. We averaged the response value over a 20-sec test period for each pre-adaptation condition. Figure 12 shows the aver-

Figure 11. Temporal change in perceived shape for conflicting-cue stimulus under the various pre-adaptation conditions shown in Figure 8. Upper panel shows the results for RVF and UVF; lower panel shows the results for KDE-P, KDE-D, and DC-S. Both panels also show the results for two control conditions (SI and KDE). The data (Subject K.U.) are smoothed temporally with a rectangular window of 3 sec.

ages with the standard error of trials. A large average value (close to that of the SI condition) implies a strong contribution of the KDE cue to the perception of the test stimulus, or a weak effect of the pre-adaptation stimulus on the adaptation of the KDE mechanism. A small average value indicates a strong contribution of the stereo cue, or a strong effect of the pre-adaptation stimulus.

Figure 12 shows that the general trends of 2 of the 4 subjects (K.U. and S.N.) are similar. The average response values for four of the conditions (UVF, RVF, KDE-P, and KDE-D) are roughly the same as for the SI condition, whereas the response values for the other two conditions (KDE and DC-S) are consistently smaller. The results for Subject H.I. also show similar tendencies, except that the value for the UVF condition seems to be somewhat small. In the results for Subject K.M., how-

ever, the average values for five of the conditions (UVF, KDE-P, KDE-D, KDE, and DC-S) seem to be smaller than for the SI condition, whereas the standard errors are fairly large. From these data, we could not determine whether this subject had an adaptation mechanism that was different from the others', or simply could not maintain a stable decision criterion across the different pre-adaptation conditions. Therefore, the following analysis is based on the results of the 3 subjects who showed similar response tendencies.

An ANOVA for the seven pre-adaptations and 3 subjects (K.U., S.N., and H.I.) revealed significant main effects of pre-adaptation [$F(6,189) = 12.93, p < .01$] and subject [$F(2,189) = 51.82, p < .01$]. The interaction between pre-adaptation and subject was not significant ($F < 1.0$). Further probing of the mean differences for

Figure 12. Response value for various pre-adaptation conditions averaged over 20 sec. Error bars show ± 1 standard error among trials. The vertical scale is adjusted to the response range of each subject.

the pre-adaptations (Tukey-Kramer's *HSD* test) indicated that there were two types of responses: a small value for KDE and DC-S and a large value for SI, UVF, RVF, KDE-P, and KDE-D.² Table 1 shows the results of the mean comparisons for all the pre-adaptation conditions (5% level). Therefore, for the 3 subjects, the average response values were significantly smaller for the KDE and DC-S conditions than for the SI, UVF, RVF, KDE-D, and KDE-P conditions.

Discussion

First consider the possibility that the self-adaptation of KDE occurs at a higher shape-reconstruction stage. If so, it is likely that the self-adaptation proceeds even when the pre-adaptation stimulus is presented at a different retinal position or is made from a velocity field of the opposite direction, as long as the pre-adaptation stimulus produces the same KDE shape. However, the average response value was not substantially smaller for the KDE-P and KDE-D conditions than for the SI condition. This suggests that only slight adaptation, if any at all, occurs at

the higher level of KDE processing, which is independent of the retinal position or velocity field. Nawrot and Blake (1991), who examined a different type of adaptation effect concerning an interaction of KDE and stereo cues, also found a retinal specificity of the adaptation effect.

Next, if the adaptation of the KDE mechanism occurs at the motion detection level, the average response value will be much lower for the UVF or RVF condition, in which a uniform or random velocity field is presented during the pre-adaptation period. However, the value for the RVF condition was roughly the same as, or only slightly lower than, that for the SI condition. As for the UVF condition, although the response values for 2 of the subjects (K.M. and H.I.) seemed to be lower, the values for the other subjects (K.U. and S.N.) were similar to those for the RVF condition. Further, the response value for the S $\rightarrow$ K+S condition in Experiment 2, in which the pre-adaptation had a uniform velocity field, was not significantly lower than the value for the no pre-adaptation (K+S) condition. Therefore, the self-adaptation to the

Table 1
Comparison of All Pairs of Average Response Values for
Various Pre-Adaptation Conditions in Experiment 3

	SI	UVF	RVF	KDE-P	KDE-D	KDE	DC-S
SI							
UVF	n.s.						
RVF	n.s.	n.s.					
KDE-P	n.s.	n.s.	n.s.				
KDE-D	n.s.	n.s.	n.s.	n.s.			
KDE	<	<	<	<	<		
DC-S	<	<	<	<	<	n.s.	

Note—Results from Tukey-Kramer's *HSD* test, using the JMP on Macintosh. Significance level is 5%.

KDE stimulus may occur at the motion detection level, but it seems difficult to account for the effect of pre-adaptation to the KDE stimulus as well as the temporal change of depth perception for the test stimulus solely by the adaptation of simple motion detectors.

Finally, let's consider whether the adaptation of the KDE mechanism occurs at the relative motion detection level. If this is the case, the average response value will be lower for the DC-S condition, in which the pre-adaptation stimulus contains relative motion components that are similar to those of the test stimulus. Consistent with this prediction, the value for the DC-S condition was as low as that for the KDE condition for all the subjects. Note that the DC-S stimulus, in which the dots moved in the direction opposite that of the dots in the test stimulus, will not enhance the adaptation at the motion detection level. Further, the pre-adaptation stimulus did not produce a percept of a vertical cylinder, thus it will not enhance the self-adaptation at the shape-reconstruction stage. Therefore, the present results support the hypothesis that the adaptation of the KDE mechanism occurs strongly at the relative motion detection level.

GENERAL DISCUSSION

To analyze the dynamics of the visual system when combining multiple depth cues, we used a stimulus that involved conflicting KDE and binocular stereopsis cues. We found a temporal change in depth perception: KDE shape perception dominates at first, but then gradually changes into stereopsis shape perception. The results of a second experiment, employing a pre-adaptation stimulus involving only a KDE cue or stereo cue, suggested that the temporal change we observed was mainly due to the self-adaptation of the KDE mechanism. In a third experiment, we tested the effects of various pre-adaptation stimuli and obtained a line of evidence supporting the idea that self-adaptation of the KDE mechanism occurs at the level of motion and relative motion detection.

Given that the relative contribution of the KDE cue gradually decreases over time due to the self-adaptation of the KDE system, there are two possible interpretations of the way the output of the KDE system changes. One is that the output of the KDE system becomes noisy and

unreliable, so less weight is given to the KDE output even though the system continues to compute the shape correctly (see Landy et al., 1991; Maloney & Landy, 1989). The other interpretation is that the output of the KDE system deteriorates, so the depth perception based on KDE diminishes over time. If the latter is the case, the shape from KDE goes flat as the KDE system adapts, regardless of the presence of other cues. To test this idea, we made an informal observation, also using a KDE stimulus simulating a view of a rotating cylinder. To remove the conflicting stereo cue, the stimulus was viewed monocularly. A KDE stimulus was presented slightly above the fixation point for about 30 sec. After a short interval, two KDE stimuli were presented, one at the same position as during the adaptation and the other below the fixation point. Even though the two stimuli were physically identical, the several people who observed the display reported that the KDE stimulus presented at the adapting position was much flatter than the comparison stimulus. A similar effect was obtained when the adapting stimulus was presented below the fixation point. The results of this observation suggest that shape perceived from KDE goes flat as the KDE system self-adapts. It is yet unknown, however, whether the flattening of the KDE shape is a necessary and sufficient condition for the decrease in the relative contribution of KDE for the conflicting-cue stimulus. It might be that the weight given to the KDE system decreases at the same time.

We found that the effect of KDE gradually decreased over time. This is contrary to the results of Hildreth, Grzywacz, Adelson, and Inada (1990) and Treue et al. (1991), which showed that the accuracy of 3-D structure recovery from the KDE stimulus improves over time. This apparent contradiction can easily be accounted for in terms of the difference in the time scale. The earlier studies were concerned with improvement in the accuracy of structure recovery over periods of up to 1 sec, whereas the temporal change we noticed here occurred at least a few seconds after the presentation of the stimulus.

Prolonged observation of an object at one depth can change the apparent distance of the object (Blakemore & Julesz, 1971) and elevate its detection threshold (Blakemore & Hague, 1972; Felton, Richards, & Smith, 1972; Stevenson, Cormack, Schor, & Tyler, 1992). These phenomena lead to an expectation that a self-adaptation of the stereopsis mechanism should also occur for our conflicting-cue stimulus, as well as for the self-adaptation of the KDE mechanism. Adaptation probably occurs in both mechanisms simultaneously, but the adaptation of the KDE mechanism is much stronger, or quicker, than that of the stereopsis mechanism. This difference in the strength of adaptation between the two mechanisms may be ecologically reasonable, because it is much rarer in everyday life to see an object rotating at the same (retinal) position for a long time than it is to fixate binocularly on a stable object at a distance.

We found that pre-adaptation to a stimulus consisting of relative motion components similar to those of the test

stimulus greatly reduced the effects of the KDE cue in the perception of the test stimulus (DC-S condition in Experiment 3). This result suggests that the KDE mechanism involves a relative motion detection stage, and that strong adaptation occurs at that stage. Although previous studies have postulated a relationship between relative motion detection and structure recovery from motion (Allman et al., 1985; Droulez & Cornilleau-Pérès, 1990; Nakayama & Loomis, 1974), we know of no direct evidence indicating that relative motion detection plays a significant role in the KDE. Our results confirm the contribution of relative motion detection to the KDE by the classical psychophysical technique of adaptation. Future research is expected to theoretically justify the role of relative motion detection in recovering structure from motion.

The present results indicate that researchers using conflicting-cue stimuli should consider the possibility of temporal change in shape perception. If there is a temporal change similar to what we observed, the interaction between multiple depth cues cannot be evaluated properly without careful control of the stimulus presentation time and the resting interval. We also demonstrated that the main cause of the observed change in depth perception is self-adaptation of the KDE system. The present results suggest that careful time control is required in studies on KDE perception, especially when the stimulus has a velocity field that continuously stimulates the same local and relative motion detectors.

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.
- ALLMAN, J., MIEZIN, F., & MCGUINNESS, E. (1985). Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area. *Perception*, 14, 105-126.
- ANDO, H. (1991). Dynamic reconstruction of 3D structure and 3D motion. *Proceedings of the IEEE Workshop on Visual Motion* (pp. 101-110). Princeton, NJ.
- BLAKEMORE, C., & HAGUE, B. (1972). Evidence for disparity detecting neurons in the human visual system. *Journal of Physiology*, 225, 437-455.
- BLAKEMORE, C., & JULESZ, B. (1971). Stereoscopic depth aftereffect produced without monocular cues. *Science*, 171, 286-288.
- BÜLTHOFF, H. H., & MALLOT, H. A. (1990). Integration of stereo, shading and texture. In A. Blake & T. Troscianko (Eds.), *AI and eye* (pp. 119-146). New York: Wiley.
- CORNILLEAU-PÉRÈS, V., & DROULEZ, J. (1989). Visual perception of surface curvature: Psychophysics of curvature detection induced by motion parallax. *Perception & Psychophysics*, 46, 351-364.
- DOSHER, B. A., LANDY, M. S., & SPERLING, G. (1989). Kinetic depth effect and optic flow: I. 3D shape from Fourier motion. *Vision Research*, 29, 1789-1813.
- DROULEZ, J., & CORNILLEAU-PÉRÈS, V. (1990). Visual perception of surface curvature: The spin variation and its physiological implications. *Biological Cybernetics*, 62, 211-224.
- FELTON, T. P., RICHARDS, W., & SMITH, R. A. (1972). Disparity processing of spatial frequencies in man. *Journal of Physiology*, 264, 621-640.
- HILDRETH, E. C., GRZYWACZ, N. M., ADELSON, E. H., & INADA, V. K. (1990). The perceptual buildup of three-dimensional structure from motion. *Perception & Psychophysics*, 48, 19-36.
- KOENDERINK, J. J., & VAN DOORN, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A*, 75, 377-385.
- LANDY, M. S., MALONEY, L. T., & YOUNG, M. J. (1991). Psychophysical estimation of the human depth combination rule. In P. S. Schenker (Ed.), *Sensor fusion III: 3-D perception and recognition* (Proceedings of the SPIE, No. 1383, pp. 247-254).
- LOOMIS, J. M., & NAKAYAMA, K. (1973). A velocity analogue of brightness contrast. *Perception*, 2, 425-428.
- MALONEY, L. T., & LANDY, M. S. (1989). A statistical framework for robust fusion of depth information. In W. A. Pearlman (Ed.), *Visual communication and image processing IV*. (Proceedings of the SPIE, No. 1199, pp. 1154-1163).
- NAKAYAMA, K., & LOOMIS, J. M. (1974). Optical velocity patterns, velocity sensitive neurons, and space perception: A hypothesis. *Perception*, 3, 63-80.
- NAWROT, M., & BLAKE, R. (1991). The interplay between stereopsis and structure from motion. *Perception & Psychophysics*, 49, 230-244.
- NORMAN, J. F., & LAPPIN, J. S. (1992). The detection of surface curvatures defined by optical motion. *Perception & Psychophysics*, 51, 386-396.
- NORMAN, J. F., TODD, J. T., TITTLE, J. S., & PEROTTI, V. J. (1993). Directional anisotropies determine the perceived structure of contradictory motion and binocular disparity fields. *Investigative Ophthalmology & Visual Science*, 34, 1082.
- REGAN, D. (1986). Visual processing of four kinds of relative motion. *Vision Research*, 26, 127-145.
- ROGERS, B. J., & COLLETT, T. S. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. *Quarterly Journal of Experimental Psychology*, 41A, 691-717.
- ROGERS, B. J., & GRAHAM, M. E. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science*, 221, 1409-1411.
- STEVENSON, S. B., CORMACK, L. K., SCHOR, C. M., & TYLER, C. W. (1992). Disparity tuning in mechanisms of human stereopsis. *Vision Research*, 32, 1685-1694.
- TREUE, S., HUSAIN, M., & ANDERSEN, R. A. (1991). Human perception of structure from motion. *Vision Research*, 31, 59-75.
- ULLMAN, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- VAN SANTEN, J. P. H., & SPERLING, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2, 300-321.

NOTES

1. If a dot appeared at the n th frame, it moved rotation angle 20° over 10 frames (unless it went behind), and reappeared at the initial position at the $(n+10)$ th frame. The start frame of each dot (n) was randomized. Thus, the dot pattern of the $(n+10)$ th frame was identical to the pattern of the n th frame. Note also that the observer did not perceive backward motion for the abrupt change of dot position, since only about 10% of the dots jumped backward for each frame. In addition, the backward-jump size was much larger than the forward-jump size, which made it difficult to find correspondence in the dense dot pattern.
2. ANOVA for the seven pre-adaptations and 4 subjects (including Subject K.M.) revealed significant main effects of pre-adaptation [$F(6,252) = 9.41, p < .01$] and subject [$F(3,252) = 27.5, p < .01$]. The interaction between pre-adaptation and subject was not significant ($F < 1.0$). Further probing of the mean differences for the pre-adaptations indicated the same results as in the main text, except for the UVF condition: There was a significant difference between the UVF and SI conditions, but no significant differences between the UVF and the other conditions.

(Manuscript received March 16, 1993;
revision accepted for publication October 6, 1993.)