

---

# Mechanisms for seeing transparency-from-motion and orientation-from-motion

---

Leo Poom

Department of Psychology, Uppsala University, Box 1225, S-751 42 Uppsala, Sweden;  
e-mail: [Leo.Poom@psyk.uu.se](mailto:Leo.Poom@psyk.uu.se)

Received 7 October 1999, in revised form 28 March 2000

---

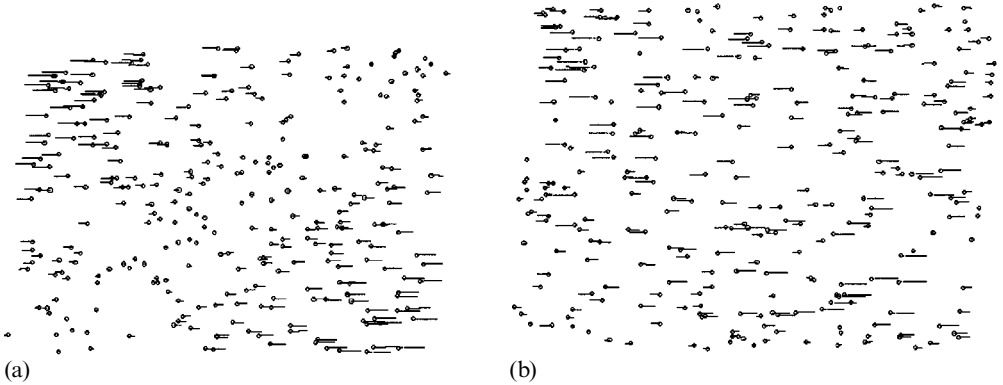
**Abstract.** Structure-from-motion (SFM) perception is hypothesised to be mediated by units that sense the near–far relationships in transparency-from-motion (TFM) and orientation-from-motion (OFM). The frequency of subjective reversals during observation of ambiguous SFM displays is considerably decreased when either the direction of rotation or the surface orientation is oscillated during the inspection. Such manipulations impede adaptation of units selectively sensitive to TFM and OFM. The results show that both OFM and TFM units are direction-selective, and that the reversal rate is unaffected by reducing TFM to zero; and support the view that depth order, when both TFM and OFM are present, is estimated by common neural units.

## 1 Introduction

The human visual system is exceptionally sensitive to information in relative motion patterns as projected on the retina (retinal flow). Since Wallach and O'Connell's (1953) first demonstrations of the kinetic depth effect, extensive research has been aimed at identifying the neural mechanisms underlying our ability to perceive three-dimensional structure-from-motion (SFM) from the two-dimensional projections in the eyes. One problem that has not received much attention is how motion information specifying perceived transparency-from-motion (TFM; overlapping motion plaids separated in depth) and perceived orientation-in-depth-from-motion (OFM) are integrated. The retinal flow patterns specifying these different components of distal structure are qualitatively different. The aim of the four experiments reported here has been to investigate how OFM mechanisms operate in the presence or absence of TFM. Are the dynamic properties the same or are they different?

Random-dot displays simulating rotating curved surfaces may cause vivid impressions of surface curvature in depth (eg Norman and Lappin 1992), although no curvature in depth can be seen when the surface is stationary. Börjesson and von Hoffsten (1973) were the first to point out that retinal flow patterns can be decomposed into basic components of relative velocity. Subsequently, Koenderink and van Doorn (1975) proved mathematically that any smooth retinal flow pattern (having a definable velocity gradient) can be decomposed into four basic types (curl, div, and two components of shear), and that these components provide information about the distal orientation of surfaces. Researchers trying to capture the process mediating surface orientation and curvature have then tried to identify the corresponding neural substrates tuned in their response patterns to these velocity gradient components (eg Reagan 1986). An illustration of smooth flow is provided in figure 1a displaying a slanted surface rotating about a vertical axis with velocity vectors depicted as line segments. As can be seen, texture elements spatially separated on the image move at different velocities, but nearby texture elements on the image stay approximately at the same relative distance.

On the other hand, perception of TFM occurs when textured, overlapping motion plaids have different velocities and are perceived as homogenous surfaces at different depths with the near surface appearing transparent between its texture elements. Such projections give rise to discontinuous retinal flow. No velocity gradients, and hence no

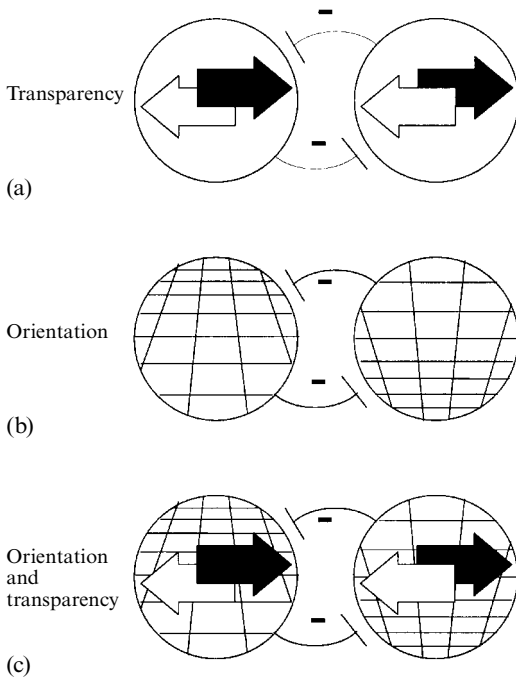


**Figure 1.** (a) An illustration of smooth flow, arising from a slanted flat rectangular surface rotating about a central vertical axis. (b) Discontinuous flow from a trapezoid 'box', with near and far surfaces separated in depth, having the same slant as the flat surface, and rotating about a central vertical axis. The front and back surfaces move with different image velocities due to the rotation. These illustrations come from the stimuli used in experiments 2 and 3.

flow components, are defined in such patterns since different velocity vectors occupy the same area on the projection plane (Andersen 1989). An example is illustrated in figure 1b, displaying a box slanted in depth and revolving about the vertical. Experimental results have provided evidence for the involvement of direction-selective and disparity-selective units in TFM perception. For example, Nawrot and Blake (1989) found that adaptation to a revolving SFM display, disambiguated by disparity information, caused a subsequent ambiguous revolving figure to be perceived as rotating in the opposite direction from that adapted to. Also, presenting frontoparallel motion in different directions in different stereoscopic depth planes can disambiguate, or prime, the direction of rotation of a sphere (Nawrot and Blake 1993). The Nawrot and Blake (1991) neural network accounts for these results, as well as the subjective reversals and depth separation in TFM patterns. In their model, direction-selective units also sensitive to the sign of binocular disparity (crossed or uncrossed) with overlapping receptive fields encode motion direction in different stereoscopic depth planes. Units sensitive to opposite directions in the same depth plane inhibit each other, whereas they excite each other if they encode motion direction in different depth planes, causing perceived depth separation even though the disparity is zero. Adaptation and reciprocal inhibitions between units tuned to opposite depth orders causes spontaneous changes of activity to take place when the information specifying the depth order is ambiguous.

The Nawrot and Blake (N&B) model is restricted to cope with depth separation in TFM, so their model needs to be extended in order to be able to explain perceived surface curvature and slant from OFM patterns. On the other hand, detectors sensitive to velocity gradients, or optic flow components (derived from smooth retinal flow), cannot account for the perception of TFM, since mathematically no velocity gradient is defined in such patterns. This problem can be solved if efficient depth separation can be made so that the gradients can be estimated separately in perceptually different depth planes.

Figure 2 summarises the main idea to be tested in this paper, namely to probe processes mediating perception of OFM and TFM. Figure 2a shows units selectively sensitive to direction and depth in TFM patterns, and the possible amount of depth separation in such patterns. Actually, the perception of surfaces is a prerequisite for the perception of transparency and, since a perceived surface always has a perceived orientation, it follows that TFM cannot be perceived without OFM. Thus, it is assumed that the TFM units shown in figure 2a are activated by frontoparallel surfaces, and hence



**Figure 2.** Retinal-flow preferences of hypothetical depth-order units. Inhibition occurs between: (a) units sensing opposite depth order in transparency-from-motion from frontoparallel motion plaids—near–far relations are coupled to left–right directions in conflicting pairs; (b) units sensing opposite sign of orientation (in three-dimensional space) from that of motion; (c) units sensitive to opposite signs of conjunctions of orientation-from-motion and transparency-from-motion.

signal such surfaces. Figure 2b shows hypothetical OFM units, selectively sensitive to the sign of slant and the possible amount of slant from motion patterns having a defined motion gradient. In situations where OFM and TFM are simultaneously present, orientation and transparency might be coded by common neural elements. Figure 2c shows the possibility that TFM and OFM are activating common units when two overlapping smooth velocity fields are observed, such that when presented together they provide discontinuous retinal flow, signaling both transparency and slanted surfaces. When motion patterns, ambiguous with respect to the depth order are presented, inhibition between units that signal opposite depth orders makes sure that only one interpretation at a time can be perceived. The dynamics of adaptation and recovery of active and inhibited units mediates the perceived depth reversals.

SFM stimuli lead to bistable perception when the depth order is ambiguous, as when parallel projection is used to display the image eg on a computer screen. There is considerable support for that adaptation of neural elements being involved in the process that causes subjective reversals during observation of ambiguous figures (eg Babich and Standing 1981; Nawrot and Blake 1989), although cognitive factors, like attention (García-Pérez 1989) and learning (Long et al 1983), influence the process. Here, a novel approach named *auto-adaptation* is used to probe the mechanisms mediating SFM perception to distinguish it from other experimental work employing adaptation. In auto-adaptation, the same ambiguous stimulus is continuously presented, representing both the adaptation and the test stimulus, while the frequency of reversals is measured. This is contrary to previous adaptation studies of SFM mechanisms where the adaptation and test stimuli differ and are presented successively (eg Nawrot and Blake 1989). Endogenous processes can be probed in isolation by using bistable stimuli; perceptual alterations due to intrinsic processes of the perceiver can then be separated from perceptual alterations due to external stimulus manipulations. The aim is to selectively impair adaptation of the hypothesised TFM and OFM units and hence reduce the frequency of subjective reversals. If the reversal rate can be manipulated according to predictions that follow from the hypothesised units, then the existence of such units

gains considerable support. Experiment 1 is a test of the N&B model in which the auto-adaptation method is used with TFM stimuli. Experiments 2 and 3 go beyond the N&B model, which is not concerned with OFM, by testing the hypothesised existence of OFM units by the method of auto-adaptation.

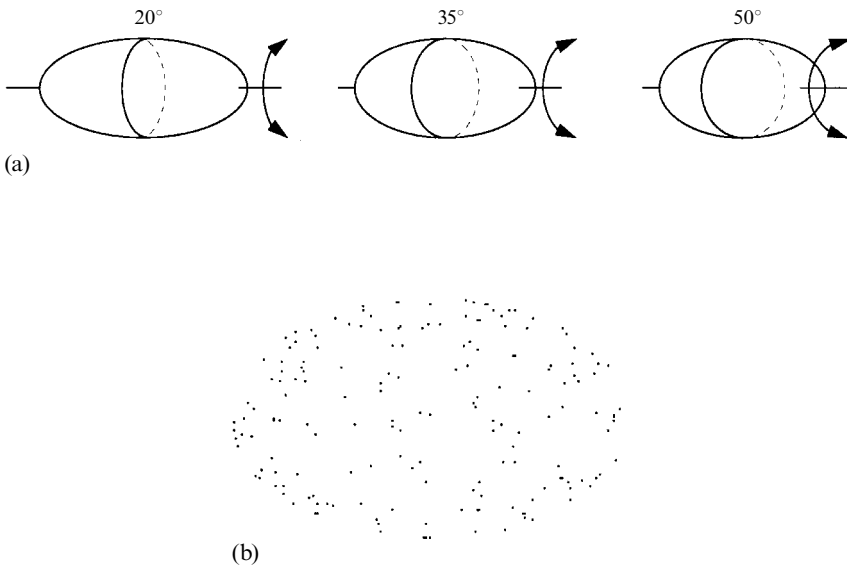
## 2 Experiment 1

It has been demonstrated earlier that the N&B model captures the relevant mechanisms that cause depth separation and accompanied depth reversals during TFM perception. The aim of experiment 1 is to further test the N&B model, to assess the auto-adaptation method, and provide some hints about the time course of adaptation and recovery. Oscillatory motion back and forth should prevent efficient adaptation of units sensitive to direction and depth because there is time for recovery in each half-period of motion. A solid shape generated by the revolution of an ellipse about its major axis (a prolate spheroid) with an oblique orientation in depth was used as stimulus to separate the detection of physical stimulus change (direction of rotation about the major axis) from subjective reversals (sign of slant of the major axis).

### 2.1 Method

**2.1.1 Participants, apparatus, and stimulus.** Three male and six female undergraduate students were paid to participate as observers. They were all naïve with respect to the purpose of the experiment.

The stimulus consisted of computer-generated parallel projections of a transparent prolate spheroid, randomly covered with 175 white dots, rotating around the principal axis against a dark background. The principal axis was horizontally oriented, but the orientation in depth (slant) of the axis was varied in different viewing conditions (figure 3a). The interstimulus interval (ISI) was always set to zero, making the stimulus onset asynchrony (SOA) equal to the presentation time of each frame. The 17-inch computer screen had a refresh rate of 75 Hz and a spatial resolution of  $1024 \times 768$  pixels. Each dot presented on the screen had the size of one pixel (about  $1.5 \text{ min of arc} \times 1.5 \text{ min of arc}$ ).



**Figure 3.** (a) Stimulus settings used in experiment 1. The prolate spheroid was rotating around the principal axis in an oscillatory fashion with different amplitudes and orientations in depth. (b) One frame of the movie sequence used in experiment 1.

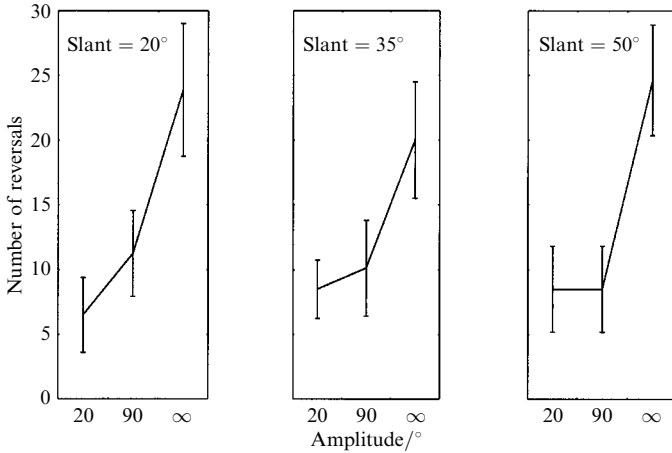
Three different orientations in depth (slants) of the prolate spheroid along with three different amplitudes of rotary oscillation around the principal axis were simulated as parallel projections on the computer screen. The slants were  $20^\circ$ ,  $35^\circ$ , and  $50^\circ$  from the frontoparallel plane. One frame in the  $35^\circ$  slant condition is shown in figure 3b. Three amplitudes of rotary oscillation were used:  $20^\circ$ ,  $90^\circ$ , and infinite. Infinite amplitude means that the prolate spheroid was simulated as revolving in the same direction during the whole inspection time. All combinations of slants and amplitudes gave a total of nine stimulus conditions. In each condition, the computer generated 360 frames of the prolate spheroid. Each frame was calculated as the parallel projection resulting from  $1^\circ$  rotation of the ellipsoid. In the  $20^\circ$  and  $90^\circ$  oscillation conditions only 20 or 90 frames of the total of 360 were used in the repeated sequences of back and forth motion. The time to complete one period of motion was 1.12 and 5 s in the oscillating conditions. When the direction of rotation was uniform, the period for a complete revolution was 10 s. The speed of rotation was  $36^\circ \text{ s}^{-1}$  in all conditions, and the ISI was set to zero, making the SOA of 28 ms equal to the presentation time of each frame. When the sequence of frames was presented in repeated succession, the impression was of a prolate spheroid smoothly revolving, or oscillating, around its major axis slanted in depth. The prolate spheroid was bistable, so that the sign of slant and near–far relations were ambiguous: either the right or the left side of it could be perceived as nearest.

*2.1.2 Design and procedure.* Each participant observed each of the nine stimulus conditions during 5 min while indicating by a key-press which side of the prolate spheroid (right or left) was perceived as nearest. The display was viewed binocularly from a distance of 90 cm. The prolate spheroid subtended  $2.5^\circ$  in the vertical direction; in the horizontal direction in the different slant conditions it subtended  $3.5^\circ$ ,  $4.1^\circ$ , and  $4.6^\circ$ . If the prolate spheroid was first perceived to be oriented so that its rightmost part was closest to the observer and a switch took place so that the leftmost part was subsequently perceived to be the closest, then the observer pressed the left-arrow key. If the impression switched from left to right then the observer pressed the right-arrow key. Between each 5 min sessions the observers took a short break while the computer created a new set of frames for the new stimulus condition.

## 2.2 Results

The results are shown after the variance between participants has been removed by subtracting, from the score of each participant, the difference between that participant's mean and the grand mean. After this operation, the mean for each participant equals the grand mean, so that the variance between participants is removed although each observer's pattern of scores remains unchanged (Loftus and Masson 1994). The reversal rate increased as the frequency of oscillatory motion decreased, and hence the amplitude increased (figure 4). The slant of the axis of rotation in depth of the prolate spheroid had no profound effect on the reversal rates. So neither the retinal size nor the orientation in depth of the axis of rotation did affect the reversal rate. The results suggest that the period of adaptation is slower than the period of recovery, since more frequent direction change decreased the reversal rate, owing to less efficient adaptation.

Summing up, the reversal rate decreases as the frequency of physical direction change increases, or, alternatively, as the amplitude of rotary oscillation decreases. Efficient adaptation of direction-and-depth-selective units is prevented if the physical direction of motion in the perceived depth planes changes during observation, and facilitated if the physical direction is the same throughout the observation. The result supports the N&B model for perceived depth separation from TFM patterns, and demonstrates that the auto-adaptation method is an efficient test of the SFM mechanisms.



**Figure 4.** The mean number of subjective reversals of a rotating prolate spheroid in 5 min of observation. The speed was held constant for different amplitudes of rotating oscillations. The number of subjective reversals increases as the period and amplitude increase, suggesting that more efficient adaptation has occurred. The vertical bars show  $\pm 1$  SE.

### 3 Experiment 2

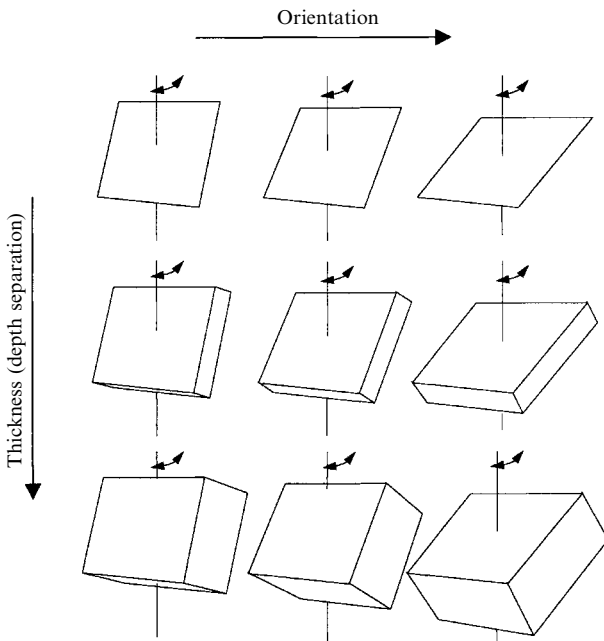
If the perceived conjunction of TFM and OFM (transparent overlapping surfaces slanted in depth) is mediated by separate units, then it seems unlikely that the reversal rates would be the same when TFM is presented in conjunction with OFM or not in conjunction with it. If separate units mediate TFM and OFM and both are activated simultaneously, then subjective reversals might occur owing to adaptation of either of these units. On the other hand, when OFM is presented alone, then only OFM units mediate the reversals. Hence, higher frequency of subjective reversals should be measured if TFM and OFM activate different neural substrates and both are present than if only OFM is present. Experiment 2 was designed to test to what extent transparency influences the reversal rate in SFM displays as compared with SFM displays with no transparency. The depth separation, or amount of discontinuous retinal flow, may be manipulated by varying the thickness of a transparent object that is revolving back and forth in an oscillatory mode. In such displays, transparency is perceived when the thickness is increased above threshold. Although it may seem reasonable, there is no evidence that physical manipulation of virtual depth separation affects the perceived transparency. Even if transparency is an all-or-none phenomenon, increasing the virtual separation in depth between the near and far surfaces increases the magnitude of discontinuous retinal flow.

Norman and Lappin (1992) reported that observers often perceived projections of frontoparallel planar surfaces rotating in depth as expanding and contracting two-dimensional patterns, not as rigid planes rotating in depth. They concluded that the presence of curvature in a direction orthogonal to the direction of rotation was necessary for the consistent perception of objects rigidly rotating in three-dimensional space. Accordingly, in preliminary observations it was noted that, when no slant along the vertical was simulated in the parallel projected image, rectangular surfaces with zero thickness (no TFM) rotating about the vertical were frequently perceived to expand and contract rather than being perceived as rigid objects. A small amount of simulated slant, such that the surface tangent made an oblique angle with the vertical axis, preserved the perceived rigidity.

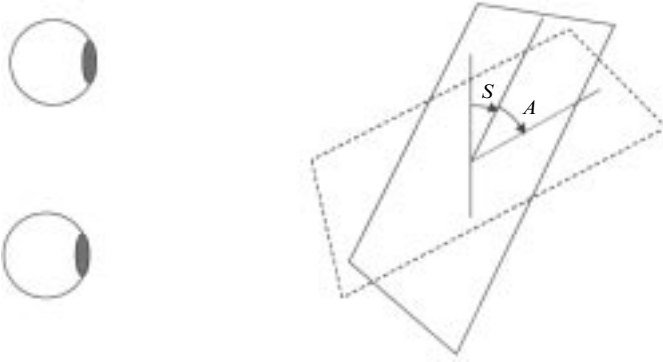
### 3.1 Method

3.1.1 *Participants, apparatus, and stimulus.* Four male and five female undergraduate students were paid to participate as observers; they were all naïve with respect to the purpose of the experiment.

The stimulus consisted of a computer-generated parallel projection of a transparent rectangular box, randomly covered with 300 white dots on a dark background. The box was simulated to rotate back and forth around a vertical axis passing through its middle section. Three different amounts of depth separation (thickness) between the near and far side of the box were used in the different conditions; 0.6 and 0.2 times the height  $h$  of the front surface of the box in two transparency conditions, and 0 depth separation in the no-transparency condition. The distal simulated rectangular surface/volume was slanted in depth:  $10^\circ$ ,  $20^\circ$ , or  $40^\circ$  along the vertical (figure 5). One frame in the  $40^\circ$  slant and  $0.6h$  thickness condition (with the velocity vectors depicted) is shown in figure 1b. No impression of depth arises from viewing such single frames. The boxes were oscillating so that the front surfaces never came closer than  $30^\circ$  from being aligned with the frontoparallel plane in the horizontal direction (angle  $S$  in figure 6). However, owing to parallel projection, the impression of which side was perceived as nearest and the near–far relationships were ambiguous as perceived on the screen. The surface was perceived to have a backward slant (topmost texture elements perceived to be at the farthest distance) when the right side of the surface was perceived to be closest. A forward slant was perceived when the left side of the surface was perceived to be closest. The box rotated  $1^\circ$  between each frame, and the period was 10 s in all conditions. Thus the angular amplitude of the motion was  $40^\circ$  (angle  $A$  in figure 6), and the speed of rotation was  $8^\circ \text{ s}^{-1}$ . The observers viewed the display binocularly from 90 cm. From that distance, the ‘box’ subtended in the vertical direction about 3 deg and in the horizontal direction between 3.2 and 4.3 deg, owing to the different slant and thickness conditions.



**Figure 5.** Illustration of the different stimuli used in experiment 2. *Rows:* slant conditions. *Columns:* depth-separation conditions. As displayed here, the right sides of the surfaces are perceived to be closest to the observer. In the actual displays, the surfaces were transparent, parallel-projected, and covered with 300 dots randomly positioned on the surface, and the perceived depth order was ambiguous.



**Figure 6.** Top view of the observer's eyes and the surface used as stimulus in experiments 2 and 3. The end-points of rotary oscillation about the vertical axis are displayed. The different amplitudes ( $A$ ) in the motion sequences started from the same deviation ( $S$ ) from the frontoparallel plane. Actually, the depth order was ambiguous, either the left or the right side of the surface/volume could be perceived as near; only one of the interpretations is shown here.

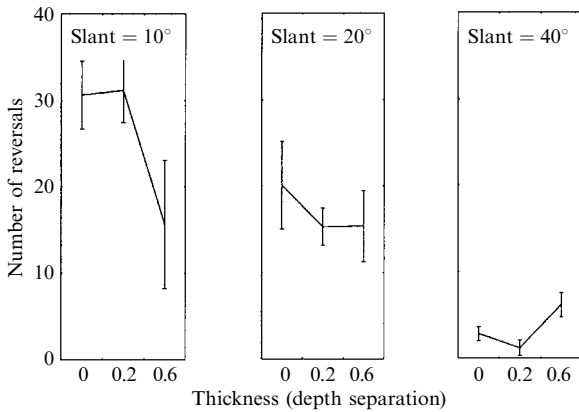
**3.1.2 Design and procedure.** The participants were asked to indicate each time the rectangular object reversed in depth by pressing the left-arrow key when the switch was from right-side to left-side nearest, and the right-arrow key when the switch was from left-side to right-side nearest. In addition, they were asked to indicate, by pressing the down-arrow key, when they were unable to see a rigid object at all.

Nine viewing conditions were created with all possible combinations of the three slant conditions and the three depth separations. The viewing conditions were randomised separately for each participant. The stimulus was demonstrated on the computer screen and a verbal instruction of the task was given after the participants had read a written instruction. The near–far relationship of the box was easily perceived: when a perceptual switch occurred it was effortlessly noted as a switch of the sign of slant accompanied with a change of the side that was perceived as closest (left or right). An observation period of 5 min was used for all stimulus conditions, so each participant made a total of 45 min of observation.

### 3.2 Results

Figure 7 shows the number of times the box was perceived to switch from one rigid interpretation to the other after the variance between observers was removed (as described in section 2.2), and whether or not a period of deformations was perceived inbetween. The switch was not counted as a real switch between rigid impressions if the same rigid impression was perceived before and after an impression of a deforming object. The number of occasions (and duration) when no rigidity was perceived was negligible. From the results in figure 7 it can be concluded that no consistent pattern of variations in reversal rate resulted from manipulation of depth separation, and hence transparency versus no-transparency conditions. Although in the  $10^\circ$  slant condition the mean number of reversals was considerably reduced when the depth separation was  $0.6h$  compared to when it was  $0.2h$  or  $0$ , the standard error (SE) is twice as large, indicating that caution should be taken before any conclusion is made based on this difference. Thus, over the different slant conditions the reversal rate has approximately the same magnitude irrespective of whether OFM is accompanied by TFM (depth separation  $0.2h$  and  $0.6h$ ) or not (depth separation  $0$ ). It should be noted that the impression of transparency was striking in both transparency conditions, although the perceived depth separation between the near and far surfaces was much less when the simulated depth separation was  $0.2h$  compared to  $0.6h$ . This might be a hint that transparency is an all-or-none phenomenon, contrary to the perceived depth separation between the near and far surfaces.





**Figure 7.** Results from experiment 2. The plots show the mean number of subjective reversals during observation of a rectangular box, oscillating about the vertical, during 5 min of observation. The depth separation between the front and back surfaces is expressed as a fraction of the height of the box. The vertical bars show  $\pm 1$  SE.

The frequency of reversals was significantly reduced when the slant of the oscillating box was increased. There was a bias in perceiving the right side of the 'box/surface' as closest (backward slant). Overall, the right side was perceived as closest twice as often or more than the left side was perceived as closest. It has been shown that observers have a perceptual bias in perceiving backward slanting surfaces rather than forward slanting ones when the sign of slant is ambiguous (Reichel and Todd 1990). Here this bias seems to increase with the slant, leading to a decreased reversal rate with increasing slant.

The main conclusion from the results of experiment 2 is that the reversal rate is largely unaffected by the presence or absence of TFM when TFM is presented in conjunction with OFM, supporting the hypothesis that when TFM and OFM are presented in conjunction they are estimated by common units.

#### 4 Experiment 3

Experiment 1 showed that units sensitive to TFM are prevented from adapting efficiently when the frequency of direction change increases (or the amplitude decreases) when a prolate spheroid rotates in depth around its main axis. Similarly, units sensitive to OFM would be expected to be prevented from adapting efficiently when the orientation of a distal surface changes during inspection, as happens when the orientations of the surfaces of a slanted box change during rotation of the box around the vertical. When the speed of rotation ( $v$ ) is constant, large amplitudes ( $A$ ) of orientation change cause more of the hypothetical OFM units to be successively activated during a period of rotary oscillation than when the amplitude is smaller. The period for unidirectional rotation is  $T = A/v$ . More time is available for these units to recover from the adaptation during periods of activity when a large number of such units share the work load successively, especially if the range of orientation selectivity for the OFM units is narrow. On the other hand, increased amplitude will be correlated with increased adaptation of the hypothetical units sensitive to TFM, since these will be subject to longer periods of stimulation. The frequency of direction change,  $f = v/A$ , can be held constant by increasing the speed of rotation as the amplitude increases so that TFM adaptation is held constant across amplitudes, and then the effects of OFM adaptation can be studied in isolation. Thus, the rate of subjective reversals should be negatively correlated with amplitude when the direction of rotation changes at a constant frequency across the amplitudes, because frequent changes of surface orientation prevent adaptation of the hypothetical units sensitive to OFM.

## 4.1 Method

**4.1.1 Participants, apparatus, and stimulus.** Eight male and fourteen female undergraduate students were paid to participate as observers. They were all naïve with respect to the purpose of the experiment.

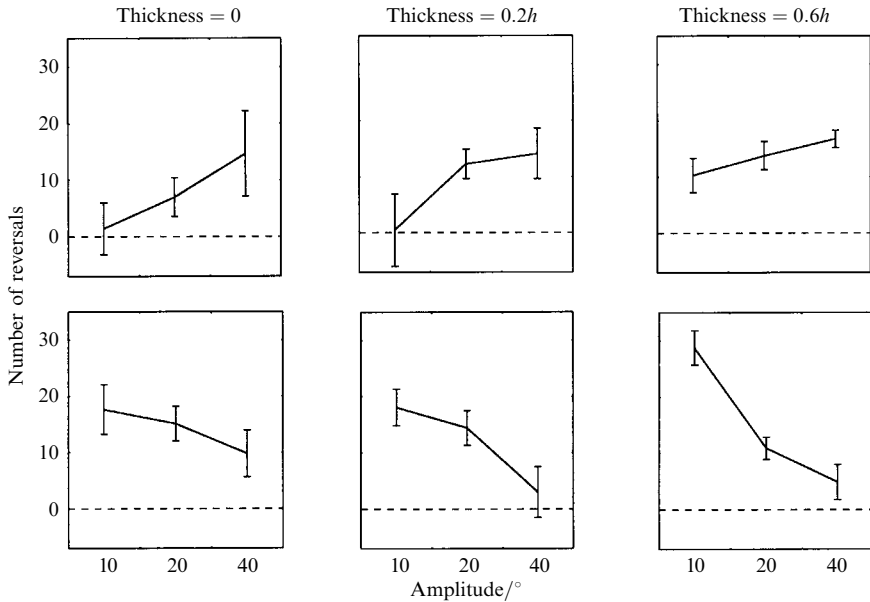
The stimulus was the same as in experiment 2, consisting of a computer-generated parallel projection of a transparent rectangular box, randomly covered with 300 white dots, displayed on a dark background. In contrast to experiment 2, the simulated slant along the vertical was held constant at  $40^\circ$  under all stimulus conditions. Two transparency conditions and one no-transparency condition were used by simulating different depth separations between the far and near side of the box (thickness conditions). The depth separations in the transparency conditions were 0.2 and 0.6 times the height,  $h$ , of the front surface of the box, and 0 in the no-transparency condition. The amplitude of rotation was also varied in three steps:  $10^\circ$ ,  $20^\circ$ , and  $40^\circ$  (angle  $A$  in figure 6). The front surface of the box never came closer to being aligned with the frontoparallel plane than  $30^\circ$  in the horizontal direction (angle  $S$  in figure 6). The vertical axis passing through the midline of the box was used as the axis of rotation. Two oscillation modes were chosen according to the relation  $T = A/v$ . Either the period  $T$  was held constant across amplitudes  $A$  by varying the speed  $v$ , or the speed was held constant for the three amplitudes. The period was 5 s in the constant-period trials, the number of frames used in the rotation sequence was 180, and the speeds of rotation were  $4^\circ \text{ s}^{-1}$ ,  $8^\circ \text{ s}^{-1}$ , and  $16^\circ \text{ s}^{-1}$  for the different amplitudes. In constant-speed conditions the numbers of frames in the sequence were 90, 180, and 360; the periods were 1.25, 2.5, and 5 s, respectively; and the speed of rotation about the vertical was  $16^\circ \text{ s}^{-1}$ .

**4.1.2 Design and procedure.** The three different amplitudes, three depth separations, and two oscillation modes were combined, resulting in a total of eighteen different stimulus conditions. The participants were divided into two groups with eleven observers each, viewing either the constant-frequency or the constant-period stimulus series. Each group was presented nine stimulus configurations. The stimulus series within each block was presented in a randomised order to each participant. Each stimulus condition was observed for 5 min. The observers' task was to indicate on the keyboard which side of the box was perceived as closest. The key-presses were made whenever the impression changed during the observation of the display. The observers were instructed to press the right-arrow key if the right side was perceived as closest and the left-arrow key if the left side was perceived as closest. If no depth was seen in the display and the dot pattern seemed to undergo deformation, then the observers were instructed to press the down-arrow key.

## 4.2 Results

As in experiment 2, deformations were ignored when counting reversals between successive rigid interpretations. The proportion of time during which observers saw the deformations was comparable to that found in experiment 2 and could be neglected. The results are shown after the variance between observers has been removed (as described in section 2.2).

The upper three panels in figure 8 display the results for the conditions where the speed of rotation was held constant across the different amplitudes. Hence, as in experiment 1, the frequency of physical change of direction varies inversely with the amplitude of rotary oscillation. As expected, the results mimic the results from experiment 1. Remarkably, the three lower panels reveal an opposite result—the reversal rate decreases with the amplitude of oscillation. The difference between the two conditions is that the frequency of direction change was kept constant in the lower panels, whereas in the top panels it decreased with amplitude. The rate of subjective reversals increased with the increase of the amplitude of oscillation (amplitude of orientation



**Figure 8.** Mean number of subjective reversals during 5 min of observation of an ambiguous rectangular box oscillating about the vertical. The depth separation between the near and far surfaces, and the amplitude of oscillation of the box were varied. *Upper panels:* The speed was held constant, so the frequency of direction change varied with amplitude, with the highest frequency at the smallest amplitude. Reversal rate increased with amplitude. *Lower panels:* The frequency of direction change was constant for the different amplitudes. Reversal rate decreased with amplitude. The vertical bars show  $\pm 1$  SE.

change) when it was accompanied by falling frequency of physical change of direction (figure 8, upper panels). However, when the frequency of physical change of direction was held constant across the different amplitudes of orientation change (lower panels in figure 8), the reversal rate decreased with increasing amplitude. Thus, adaptation of the hypothetical units selective to OFM determines the reversal rates when the frequency of direction change is held constant, so that the units selective to TFM are equally adapted at all amplitudes of orientation change.

The result in the no-transparency conditions, presented in the leftmost panels of figure 8, is especially interesting. When the frequency of physical change of direction increases (alternatively, when the amplitude decreases), the reversal rate decreases even though no transparency is present (figure 8, upper leftmost panel). When the frequency is held constant, there is an inverse relation between reversal rate and amplitude (figure 8, lower leftmost panel). These results indicate that units sensing OFM alone are direction-selective since frequent direction changes hamper efficient adaptation when only OFM is present.

#### 5 Experiment 4

Since speed co-varied with the amplitude of rotation in the constant-period conditions in experiment 3, and amplitude co-varied with the period of rotation in experiment 1, nine additional observers were tested in a control experiment. The stimulus consisted of computer-generated parallel projections of transparent spheres, randomly covered with 50 dots, subtending 3.2 deg, and rotating continuously in the same direction around the vertical. The technical and methodological details were the same as in the previous experiments. The speed of rotation was used as the independent variable and the period of one revolution was 20, 10, or 5 s in the different conditions.

### 5.1 Results

The mean of the number of reversals and SE over all observers during 15 min of observation were  $32 \pm 4$ ,  $30 \pm 3$ , and  $36 \pm 7$ , for the 20, 10, and 5 s periods of rotation. Thus, it seems unlikely that the varying speed of texture elements accounts for the data in experiment 3, where the speed was varied to keep the frequency of direction change constant across the different amplitudes of rotation. Also, since reversals occurred about every 30 s in all speed conditions, the amplitude of perceived unidirectional rotation varied between about 1.5 to 6 complete revolutions between successive reversals. Thus, it seems unlikely that the increased amplitude of rotation, rather than the increased period, caused the increased rate of reversals in experiment 1.

## 6 Discussion

The results from experiments 1 and 4 support the N&B model of TFM perception by demonstrating that the reversal rate decreases when the revolving shape frequently changes the direction of its rotation during observation. The direction change disrupts the activity of direction-selective units and prevents them from adapting efficiently. Experiment 2 went beyond the N&B model in showing that the reversal rate was largely unaffected on manipulating the amount of TFM. Surfaces with ambiguous signs of slant caused subjective reversals equally frequently irrespective of whether they were depicted as 'boxes' with their front and back surfaces separated in depth (both TFM and OFM present) or if the depth separation was zero (only OFM present). This result supports the idea that information about OFM and TFM is picked up by common neural elements, because it is unlikely that the reversal rate would be unaffected if the number of sites that mediate the reversals differed between conditions. Experiment 3 together with experiment 4 demonstrated that direction-selective mechanisms are present when TFM and OFM are presented in conjunction or when only OFM is present, and that the reversal rate can be decreased by increasing the amplitude of orientation change with the frequency of direction change kept constant across conditions. When the surface orientation changes during inspection, efficient adaptation of units selective to OFM is prevented owing to the frequent disruptions of the activity of these units. These similarities between OFM and TFM further support the idea that TFM and OFM are mediated by a common neural substrate. Some specialised feature analysers signal TFM when the orientation is frontoparallel (as shown schematically in figure 2a), others signal surface slant, or OFM, of nontransparent surfaces (figure 2b), still others signal combinations of TFM and OFM (figure 2c).

The present results can be incorporated into a model where units sensitive to smooth retinal velocity gradients in specific depth planes mutually inhibit each other if their gradient preferences differ and they signal the same depth, but excite each other if they signal different depth planes. In this proposal, it is assumed that the velocity gradients make sense ecologically in that they bear information about OFM. As such, they do not need to conform to any of the basic components in the formal mathematical descriptions by Koenderink and van Doorn (1975), but rather represent combinations of eg shear and div (Irvin et al 1999). The interactions between such units push discontinuous velocity fields apart so that OFM can be estimated separately at different depths, or binocular disparity may be used to accomplish the same task. A smooth velocity gradient specifying an ambiguous opaque surface orientation may activate two different sets of units that selectively signal the opposite signs of possible surface orientations from the ambiguous display. Only one set of these units is simultaneously active owing to inhibitory connections between sets sensing opposite signs of slant, so that only one surface is seen at a time.

There is evidence that the units selective to OFM/TFM and their mutual interactions might be found in area MT, although interactions with other areas, such as V3 and

V4 might be important for the perception of dynamic form (Vaina et al 1990; Zeki 1993). Correlation between neural firing in area V5 (MT) and the perceived direction of rotation in depth of bistable SFM stimuli has been found in monkeys; these cells also show disambiguous firing when the SFM perception is disambiguated by disparity in the display (Bradley et al 1998). V5 neurons also show sensitivity for the orientation in depth for nontransparent surfaces specified by velocity gradients (Treue and Andersen 1996; Xiao et al 1997).

Other evidence for the existence of units selectively sensitive to OFM comes from studying the aftereffects of surface slant. An aftereffect of slant with opposite sign is perceived when viewing a frontoparallel test plate immediately after the adaptation to a surface slanted in depth (Bergman and Gibson 1959; Epstein and Morgan-Paap 1974; Gibson and Radner 1937; Graham and Rogers 1982). Slant aftereffects are perceived whether the same or different information-carrying channels are used (eg texture, disparity, and motion) in the adaptation and test conditions (Balch et al 1977; Poom and Börjesson 1999). These results indicate the existence of modal independent neural units sensitive to the orientation of surfaces in three-dimensional space, so that different information-carrying media are integrated in the neural substrate before the estimation of structure is performed. It might be that the same process causes these aftereffects of surface orientation and the subjective reversals in SFM when the surface orientation is ambiguous. If so, the neural substrate that causes the perceived depth reversals when viewing stationary bistable perspective displays might be the same as the substrate causing reversals in TFM and OFM, although the information-bearing-media differ.

**Acknowledgements.** This research was partly supported by grants from the Faculty of Social Sciences, Uppsala University. I am grateful to Erik Börjesson who commented upon an earlier draft of this article, and for the valuable advice given by Mike Harris and another anonymous reviewer.

## References

- Andersen G J, 1989 "Perception of three-dimensional structure from optic flow without locally smooth velocity" *Journal of Experimental Psychology: Human Perception and Performance* **15** 263–371
- Babich S, Standing L, 1981 "Satiation effects with reversible figures" *Perceptual and Motor Skills* **52** 203–210
- Balch W, Milewski A, Yonas A, 1977 "Mechanisms underlying the slant aftereffect" *Perception & Psychophysics* **21** 581–585
- Bergman R, Gibson J J, 1959 "The negative after-effect of a surface slanted in the third dimension" *American Journal of Psychology* **72** 364–374
- Bradley D C, Chang G C, Andersen R A, 1998 "Encoding of three-dimensional structure-from-motion by primate area MT neurons" *Nature (London)* **392** 714–717
- Börjesson E, Hoffsten C von, 1973 "Visual perception of motion in depth: Application of a vector model to three-dot motion patterns" *Perception & Psychophysics* **13** 169–179
- Epstein W, Morgan-Paap C L, 1974 "Aftereffect of inspection of a perspectival stimulus for slant in depth: A new normalization effect" *Perception & Psychophysics* **16** 299–302
- García-Pérez M A, 1989 "Visual inhomogeneity and eye movements in multistable perception" *Perception & Psychophysics* **46** 397–400
- Gibson J J, Radner M, 1937 "Adaptation after-effect and the contrast in the perception of tilted lines. I. Quantitative studies" *Journal of Experimental Psychology* **20** 453–467
- Graham M, Rogers B, 1982 "Simultaneous and successive contrast effects in the perception of depth from motion-parallax and stereoscopic information" *Perception* **11** 247–262
- Irvins J, Porrill J, Frisby R, Orban G, 1999 "The 'ecological' probability density function for linear optic flow: Implications for neurophysiology" *Perception* **28** 17–32
- Koenderink J J, Doorn A J van, 1975 "Invariant properties of the motion parallax field due to movement of rigid bodies relative to an observer" *Optica Acta* **22** 773–791
- Loftus G R, Masson M E J, 1994 "Using confidence intervals in within-subject designs" *Psychonomic Bulletin & Review* **1** 476–490
- Long G M, Toppino T C, Kostenbauder J F, 1983 "As the cube turns: Evidence for two processes in the perception of a dynamic reversible figure" *Perception & Psychophysics* **34** 29–38

- 
- Nawrot M, Blake R, 1989 "Neural integration of information specifying structure from stereopsis and motion" *Science* **244** 716–718
- Nawrot M, Blake R, 1991 "A neural network model of kinetic depth" *Visual Neuroscience* **6** 219–227
- Nawrot M, Blake R, 1993 "On the perceptual identity of dynamic stereopsis and kinetic depth" *Vision Research* **33** 1561–1571
- Norman J F, Lappin J S, 1992 "The detection of surface curvatures defined by optical motion" *Perception & Psychophysics* **51** 386–396
- Poom L, Börjesson E, 1999 "Perceptual depth synthesis in the visual system as revealed by selective adaptation" *Journal of Experimental Psychology: Human Perception and Performance* **25** 504–517
- Reagan D, 1986 "Visual processing of four kinds of relative motion" *Vision Research* **26** 127–145
- Reichel D R, Todd J T, 1990 "Perceived depth inversion of smoothly curved surfaces due to image orientation" *Journal of Experimental Psychology: Human Perception and Performance* **16** 653–664
- Treue S, Andersen R A, 1996 "Neural responses to velocity gradients in macaque cortical area MT" *Visual Neuroscience* **13** 797–804
- Vaina L M, Lemay M, Bienfang D C, Choi A Y, Nakayama K, 1990 "Intact 'biological motion' and 'structure from motion' perception in a blind subject with impaired motion mechanisms: a case study" *Visual Neuroscience* **5** 353–369
- Wallach H, O'Connell D N, 1953 "The kinetic depth effect" *Journal of Experimental Psychology* **45** 205–217
- Xiao D K, Marcar S E, Raiguel G A, Organ G A, 1997 "Selectivity of macaque MT/V5 neurons for surface orientation in depth specified by motion" *European Journal of Neuroscience* **9** 956–964
- Zeki S, 1993 *A Vision of the Brain* (Cambridge: Blackwell Science)