PERGAMON

Vision Research 39 (1999) 881–886

Vision Research

Rapid communication

Feature specific segmentation in perceived structure-from-motion

Hyung-Chul O. Li *, Frederick A. A. Kingdom

McGill Vision Research Unit, Department of Ophthalmology, McGill University, 687 Pine Avenue West, Rm H4-14, Montréal, Québec, H3A 1A1, Canada

Received 2 July 1998; accepted 7 September 1998

Abstract

Motion information is important to vision for extracting the 3-D (three-dimensional) structure of an object, as evidenced by the compelling percept of three-dimensionality attainable in displays which are purely motion-defined. It has recently been shown that when subjects view a rotating transparent cylinder of dots simulated with parallel projection, they rarely perceive rotation reversals which are physically introduced (Treue, Andersen, Ando & Hildreth, *Vision Research*, *35*;1995:139–148). We show however that when the elements defining the cylinder are oriented, the number of perceived reversals increases systematically to near maximum as the difference between element orientations on the two surfaces increases. These results imply that structure-from-motion mechanisms are capable of exploiting local feature differences between the different surfaces of a moving object.

Keywords: Segmentation; Structure-from-motion; Motion transparency

1. Introduction

Motion information is important to vision for extracting the 3-D (three-dimensional) structure of an object, as evidenced by the compelling percept of depth in displays which simulate the motion properties of 3-D objects but in which all other depth cues, such as stereopsis and perspective, have been eliminated (Wallach & O'Connell, 1953; Braunstein, 1962; Rogers & Graham, 1979; Todd, 1984; Ullman, 1984). A stimulus that has recently become popular in studying structure-from-motion (SFM) processing is a rotating cylinder (or sphere) of randomly positioned dots simulated by parallel (orthographic) projection (Braunstein, Andersen & Riefer, 1982; Andersen & Braunstein, 1983; Nawrot & Blake, 1989, 1991; Treue, Husain & Andersen, 1991; Treue, Andersen, Ando & Hildreth, 1995; Jiang, Pantle & Mark, 1998). With parallel projection the front and back surfaces of the rotating cylinder are ambiguous and hence also is the direction of rotation, yet subjects perceive a compelling impression of a rigid three-dimensional object rotating in one direction or the other. Often observers experience spontaneous rotation direction reversals (Nawrot & Blake, 1989, 1991). Treue et al. (1995) and Li (1996) have recently shown that in addition, observers typically *miss* reversals when they are physically introduced. In Treue et al.'s study, even when observers were requested to track a specific surface with the help of a distinctive dot (e.g. an enlarged dot), they rarely perceived reversals, instead perceiving the dot to switch from the front to the back surface or vice versa. Treue et al. suggested that the lack of perceived rotation reversals was a consequence of the SFM object being represented as a 3-D surface rather than as a group of dots with a particular arrangement. They argued that local features are not explicitly represented in the surfaces of a SFM object, resulting in observers rarely perceiving the reversal of the rotation of the overall cylinder even when all the local dots reversed their directions.

* Corresponding author. Fax: +1-514-843-1691; e-mail: hyung@jiffy.vision.mcgill.ca.

PII: S0042-6989(98)00247-8

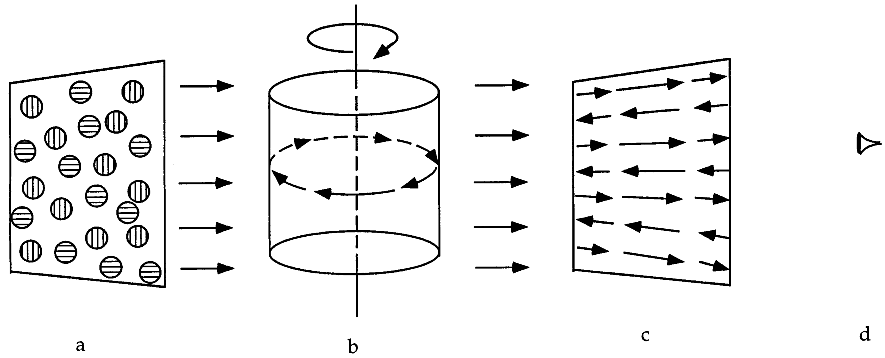

Fig. 1. Diagrammatic representation of how the rotating cylinder was generated. (a) Gabor micropatterns are randomly plotted on a 2-D square, and (b) projected onto a rotating cylinder with two transparent surfaces; (c) shows the resulting stimulus as seen by the observer (d).

While Treue et al. (1995) showed that the particular *arrangement* of dots on the cylinder does not appear to be explicitly represented, the question remains as to whether the *type* of dot nevertheless is. What if the dots on each surface were different along some dimension, such as colour, orientation or size? Would the visual system ignore such differences, with the consequence that reversals would continue to be missed, or would the visual system label the two surfaces according to dot type, perhaps enabling the reversals to be perceived? We decided to address this question using a rotating cylinder constructed from Gabor micropatterns with different orientations. The basic stimulus arrangement is shown in Fig. 1, with the two main conditions illustrated in Fig. 2. Fig. 2a illustrates the unsegregated condition, in which the front and back surfaces of the simulated cylinder contained both micropattern orientations in equal numbers. Fig. 2b illustrates the segregated condition, in which the front and back surfaces contained different micropattern orientations. The difference in orientation between the two types of micropattern, whether segregated by surface or not, was the main independent variable. With three test subjects, two of whom were naive as to the purpose of the experiment, we compared the proportion of perceived reversals of the cylinder between the segregated and unsegregated conditions.

2. Methods

2.1. Stimuli

2.1.1. Display generation

The stimuli were generated by a PowerMac 8500/180 with 8 bits-per-gun intensity resolution, and displayed on a 17″ NEC MultiSync XV17 + RGB video monitor (640H X 480V pixel resolution; P22 phosphors; 120 Hz frame rate). The screen non-linearity was gamma-corrected following calibration of the three gun luminances with a Universal photometer (Optikon).

2.1.2. Gabor micropatterns

The Gabor micropatterns were generated using the function:

$$L(x,y) = M + A \times \exp[-(x^2+y^2)/(2\sigma^2)] \times \sin\{2\pi f[x^2\cos(\theta)+y^2\sin(\theta)]\}$$

where M was mean luminance, A amplitude of 50%, σ the space constant of 0.076°, f the spatial frequency of 0.44 cycles/deg, and θ the orientation of the carrier which varied from 0 to 90°. The function was clipped at a diameter of 0.34 deg. The phase of the Gabor was set to make it odd-symmetric, ensuring that its mean luminance was the same as that of the background. When two Gabor micropatterns overlapped, their amplitudes but not DC levels were added.

2.1.3. Simulated rotating cylinder

A transparent rotating cylinder was simulated using parallel projection, that is without perspective cues. Fig. 1 illustrates the stimulus construction. The radius of the cylinder was 3.4° and its height was 6.8°. The rotation speed was 3°/frame (90°/s), and the number of Gabor micropatterns presented in each frame was 400 (200 for each surface). The life-time of each micropattern was infinite, and in order to simulate rotation, the micropatterns moved more slowly at the edges of the cylinder compared to its centre. Thus the density of micropatterns was slightly higher at the edges. The stimulus comprised 400 motion frames, with each motion frame being repeated for four monitor frames (33.3 ms each). Total stimulus presentation time was 13.32 s. Eight rotation reversals were physically introduced during each stimulus presentation. The time of each reversal was randomized with the following two constraints: (1) The first and the last reversals were not introduced during the first and last second; (2) the interval between consecutive reversals was a minimum of 500 ms.

The transparent surfaces of the cylinder could be either segregated or unsegregated in terms of the orien-

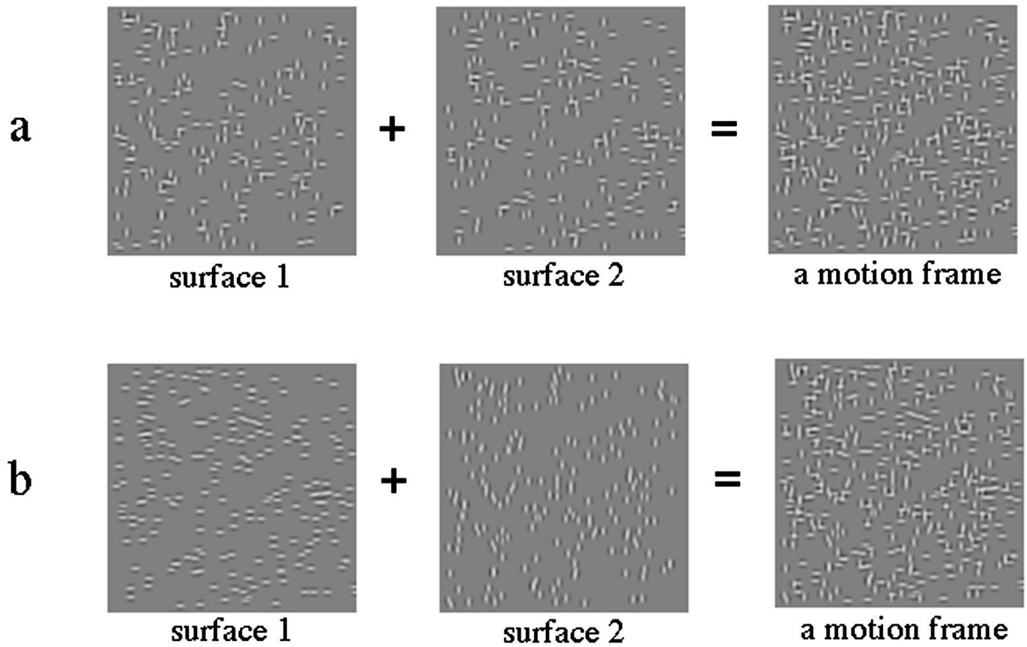

Fig. 2. 2-D view of the 0°-plus-90° (vertical-plus-horizontal) micropattern conditions. (a) Unsegregated condition in which the two transparent surfaces were not differentiated by micropattern orientation. (b) Segregated condition in which the two transparent surfaces were differentiated by micropattern orientation. There was no difference between the stationary frames of the two conditions.

tation of the Gabor micropatterns. In the unsegregated condition, both surfaces contained both orientations, either 0° (vertical) and 90°, 45° and 90°, 60° and 90°, 70° and 90°, or 90° and 90°. Fig. 2a shows the unsegregated 0° and 90° condition. In the segregated condition, one surface comprised micropatterns with an orientation of 90°, and the other surface micropatterns of one of the four orientations: 0°, 45°, 60° or 70°. Fig. 2b shows the arrangement for the segregated 0° and 90° condition.

2.2. Procedure

Before each experiment, subjects were given enough practice to familiarize themselves with the task and achieve near-asymptotic levels of performance. Before each session, subjects adapted to a blank gray screen for 1 min. On every trial, a stimulus randomly chosen from one of the segregated or unsegregated stimuli sets was presented, and subjects were instructed to press a key every time they perceived the direction of rotation to reverse. The subject observed the stimulus with his/her dominant eye, the non-dominant eye being occluded. The inter-trial interval was about 5 s. Each stimulus was presented three times in random order during each session, and each session was repeated three times, making a total of 72 physical rotation reversals per condition. The distance between the monitor screen and the observers was 57 cm. The proportion of perceived reversals across the three sessions was calculated for each condition.

3. Results

The results are shown in Fig. 3, which plots the proportion of perceived reversals as a function of the orientation difference between the micropatterns in the stimulus. The two curves are for the segregated (open symbols) and unsegregated (closed symbols) conditions. For the unsegregated conditions, subjects perceived only about 5–25% of the reversals on average, and the perceived reversal rate did not vary with the difference in micropattern orientation. In the segregated condition, on the other hand, the perceived reversal rate rose systematically to near 100% as the difference in micropattern orientation between the two surfaces increased to its maximum at 90°. When the data were collapsed across the three subjects, the interaction between condition (segregated vs unsegregated) and orientation difference was highly significant ($F(4, 8) = 22.552$, $P < 0.0005$). A post hoc analysis showed that the difference between the two conditions was statistically significant at the $P = 0.01$ level for all orientation differences except the 0° difference condition.

To demonstrate that our findings are not simply reducible to the ability of subjects to detect the orientation differences between the two surfaces of the cylinder, subject HL performed the following control experiment. Using the same set of rotating cylinders as employed in the first experiment, HL was required to decide on each trial simply whether the surfaces of the cylinder were segregated or not by micropattern orien-

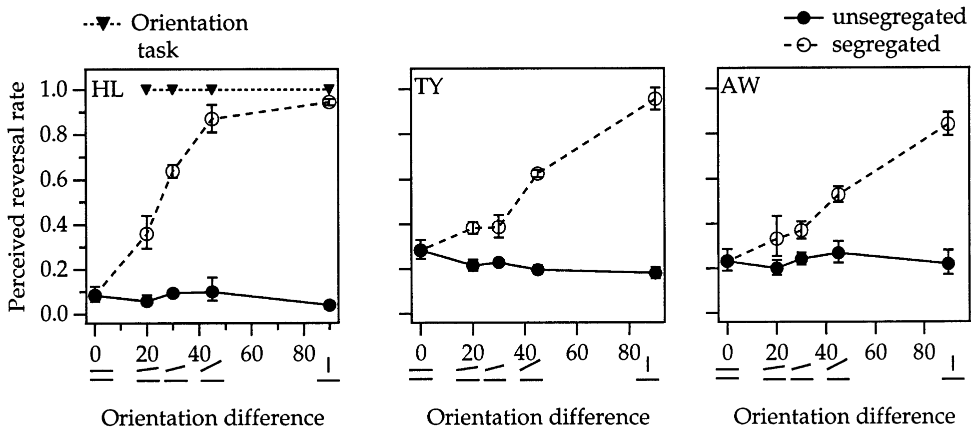

Fig. 3. The perceived reversal rate as a function of the difference in micropattern orientation of the two component surfaces of the cylinder. One orientation was fixed at 90°, and the other varied from 90 to 0°. Open circles show the segregated condition, filled circles the unsegregated condition. The filled triangles in HL's graph indicate the proportion correct for discriminating whether the cylinder was segregated or unsegregated. The error bars are standard errors calculated from the three repetitions on each condition.

tation. The stimuli were presented as forced-choice pairs for 1 s each. The results are shown as the triangles in HL's data in Fig. 3. HL's performance was 100% correct in all conditions. This clearly demonstrates that perceiving direction reversals involves mechanisms beyond those used simply for detecting orientation differences.

It would appear that the higher perceived reversal rates in the segregated conditions are not due to the disruption of the 3-D percept, as all subjects reported that the percept of a rigid rotating cylinder was equally compelling in both segregated and unsegregated conditions. A potential problem nevertheless with these results concerns the choice of micropattern orientations that we employed. In the stimuli producing the largest difference in perceived reversal rates between the segregated and unsegregated conditions, the micropatterns were vertical (0°) and horizontal (90°). As the 2-D direction of motion in the experiment was always horizontal, it is possible that the different micropattern orientations might have produced different motion strengths, providing a spurious segregation cue. To test this, we compared perceived reversal rates for segregated and unsegregated stimuli made up from two oblique micropattern orientations (+45 and −45° oblique to vertical). The results are shown in Fig. 4, which also shows perceived reversal rates for stimuli made from each of the two oblique micropattern orientations presented alone. Although the difference in the perceived reversal rates between the segregated and unsegregated conditions was slightly smaller compared to when the micropattern orientations were vertical and horizontal, it is still substantial[1], demonstrating that the effect is not specific to any particular orientation pair. When the data were collapsed over three subjects, the difference in perceived reversal rates was again highly significant ($F(3, 6) = 64.629$, $P < 0.0001$). The difference between the segregation condition and each of three unsegregated conditions, was also statistically significant at the $P = 0.05$ level, whereas no significant difference was found between any pair of the unsegregated conditions.

The previous results were obtained using infinite lifetime micropatterns. In Treue et al.'s (1995) study, the dot lifetime was limited to 200 ms. Would we obtain similar results using limited lifetime Gabor micropatterns? To test this we compared perceived reversal rates for segregated and unsegregated stimuli made up from two oblique micropattern orientations (+45 and −45° oblique to vertical), using micropatterns with 200 ms lifetimes. The results are shown in Fig. 5. Although the rate of perceived rotation reversals was in general less compared to when infinite lifetime mi-

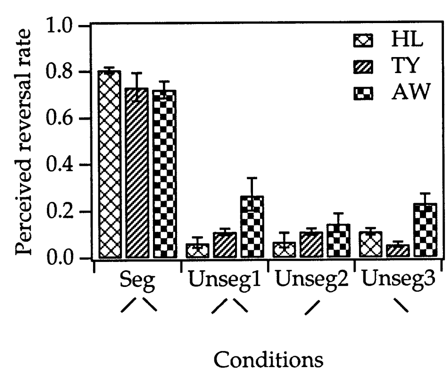

Fig. 4. Perceived reversal rates using +45° and −45° Gabor orientations. Seg = segregated condition. Three unsegregated conditions were tested: Unseg1 = both micropatterns in equal amounts; Unseg2 = +45° only; Unseg3 = −45° only.

[1] We used vertical and horizontal orientations for the main experiment because with oblique orientations the rotation direction is perceived as deviated from the left or rightward motion, and according to subjects' subjective reports this results in a weaker percept of 3-D depth.

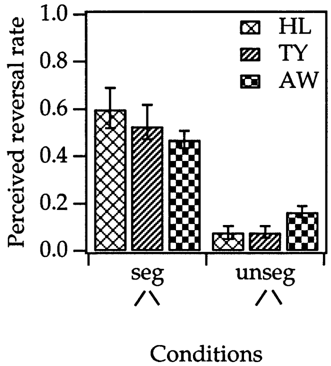

Fig. 5. Perceived reversal rates using +45° and −45° micropattern orientations using limited lifetime Gabors.

cropatterns were employed, observers still perceived significantly more rotation reversals in the segregated compared to unsegregated conditions (across subjects: $F(1, 2) = 45.28$, $P < 0.05$).

4. Discussion

These results imply that SFM and/or motion transparency mechanisms are sensitive to differences in the local feature content of object surfaces, even though insensitive to the particular arrangement of those features. Treue et al. (1995) and Hildreth, Ando, Andersen and Treue (1995) have emphasized the role of surface interpolation in the processing of SFM. According to them, an interpolation algorithm is employed to derive a complete representation of a surface from sparse depth information. In particular, Hildreth et al. (1995) suggested that local features are grouped by 2-D direction and speed prior to surface reconstruction, with interpolation performed separately on each group. Our results imply that the *type* of feature could therefore also be an important grouping factor in the derivation of the component surfaces of SFM objects. However, it is not necessarily the case that different feature types are grouped *prior* to the derivation of the surfaces of an SFM object. Our results could just as easily be interpretable in terms of a mechanism which labels the surfaces of an SFM object by feature type *after* those surfaces have been derived using the purely motion properties of the stimulus. Further experiments will need to be carried out to distinguish between a feature-grouping and feature-labeling explanation of the results of this study.

Treue et al. (1995) described how an enlarged dot appeared to jump from the foreground to the background of a rotating cylinder in cases where an introduced rotation reversal was not perceived. In our experiment, when an introduced rotation reversal was not perceived, observers reported seeing the stimulus jolt with a kind of random motion seen on both surfaces. They did not however report seeing the surfaces swap in depth. In the segregated conditions, on those occasions when observers also did not perceive an introduced rotation reversal, they reported seeing the features on the surface change (e.g. from vertical to horizontal or vice versa), but in most instances again did not see a depth reversal.

There have been a number of studies showing that additional depth information such as occlusion and disparity disambiguate the rotation direction of spheres simulated with parallel projection (Braunstein et al., 1982; Andersen & Braunstein, 1983; Braunstein, Andersen, Rouse & Tittle, 1986) and Necker cube (Dosher, Sperling & Worst, 1986). The differences in micropattern orientation between the surfaces of our segregated-by-feature rotating cylinders do not however provide any depth information. As a consequence the direction of rotation could still be either clockwise or anti-clockwise at the start of the stimulus presentation. In spite of the arbitrariness of the initial direction of rotation however, subjects in the segregated condition still perceived clear reversals of rotation direction when they were physically introduced. The visual system appears to set up an initial hypothesis of rotation direction, which becomes the basis for comparing any subsequent physical changes to the stimulus.

A number of studies have shown that the perception of coherent motion of plaids consisting of two orthogonal sine wave gratings decreases as the component gratings differ along some dimension, and these include: colour (Kooi, DeValois, Grosof & Switkes, 1989; Krauskopf & Farell, 1990), relative contrast and spatial frequency (Movshon, Adelson, Gizzi & Newsome, 1985), perceptual transparency (Stoner, Albright & Ramachandran, 1990) and binocular disparity (Adelson, 1984). There are many differences between the plaid stimulus/task and our rotating cylinder/task, and so it would be premature to conclude that the mechanisms involved are the same. Nevertheless, it will be interesting to discover whether similar rules for feature specificity determine the perceived reversal rates for the simulated rotating cylinder as determine the perceived coherence of plaids.

5. Conclusion

We have shown that an orientation difference between the micropatterns on the two surfaces of a simulated rotating cylinder dramatically affects its perceived properties. The method we have used to reveal feature-specific segmentation in SFM processing can be easily applied to other types of feature difference, e.g. colour, luminance polarity, local spatial frequency, as well as higher-order features, such as particular feature configurations. The technique described here can be used to understand more fully the characteristics of the mechanisms responsible for SFM.

Acknowledgements

This research was supported by Medical Research Council of Canada grant (Grant No. MT 11554) awarded to Fred Kingdom.

References

Adelson, E. H. (1984). Binocular disparity and the computation of two-dimensional motion. *Journal of the Optical Society of America*, *A1*, 1266.

Andersen, G. J., & Braunstein, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, *34*, 356–362.

Braunstein, M. L. (1962). Depth perception in rotating dot patterns: effects of numerosity and perspective. *Journal of Experimental Psychology*, *64*, 415–420.

Braunstein, M. L., Andersen, G. J., & Riefer, D. M. (1982). The use of occlusion to resolve ambiguity in parallel projections. *Perception & Psychophysics*, *31*, 261–267.

Braunstein, M. L., Andersen, G. J., Rouse, M. W., & Tittle, J. S. (1986). Recovering viewer-centered depth from disparity, occlusion, and velocity gradients. *Perception & Psychophysics*, *40*, 216–224.

Dosher, B. A., Sperling, G., & Worst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research*, *26*, 973–990.

Hildreth, E. C., Ando, H., Andersen, R. A., & Treue, S. (1995). Recovering three-dimensional structure from motion with surface reconstruction. *Vision Research*, *35*, 117–137.

Jiang, Y., Pantle, A. J., & Mark, L. S. (1998). Visual inertia of rotating 3-D objects. *Perception & Psychophysics*, *60*, 275–286.

Kooi, F. L., DeValois, K. K., Grosof, D. H., & Switkes, E. (1989). Coherence properties of colored moving plaids. *Investigative Ophthalmology and Visual Science* (*Suppl.*), *30*, 389.

Krauskopf, J., & Farell, B. (1990). Influence of colour on the perception of coherent motion. *Nature*, *348*, 328–331.

Li, H.-C. O. (1996). *Direction perception of rotating 3-D cylinders*. Doctoral dissertation. University of Wisconsin, Madison.

Movshon, J. A., Adelson, E. A., Gizzi, M., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass & C. G. Cross, *Study group on pattern recognition mechanisms*. Vatican City: Pontifica Academia Scientarum.

Nawrot, M., & Blake, R. (1989). Neural integration of information specifying structure from stereopsis and motion. *Science*, *244*, 716–718.

Nawrot, M., & Blake, R. (1991). The interplay between stereopsis and structure from motion. *Perception & Psychophysics*, *49*, 230–244.

Rogers, B. J., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*, 125–134.

Stoner, G. R., Albright, T. D., & Ramachandran, V. S. (1990). Transparency and coherence in human motion perception. *Nature*, *344*, 153–155.

Todd, J. T. (1984). The perception of three-dimensional structure from rigid and nonrigid motion. *Perception & Psychophysics*, *50*, 509–523.

Treue, S., Andersen, R. A., Ando, H., & Hildreth, E. C. (1995). Structure-from-motion: Perceptual evidence for surface interpolation. *Vision Research*, *35*, 139–148.

Treue, S., Husain, M., & Andersen, R. (1991). Human perception of structure from motion. *Vision Research*, *31*, 59–75.

Ullman, S. (1984). Maximizing rigidity: the incremental recovery of 3-D structure from rigid and nonrigid motion. *Perception*, *13*, 255–274.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*, 205–217.