Perception, 2006, volume 35, pages 993 – 997

DOI:10.1068/p5568

Last but not least

The bar – cross – ellipse illusion: Alternating percepts of rigid and nonrigid motion based on contour ownership and trackable feature assignment

Abstract. We present a new multistable stimulus generated by continuously rotating an ellipse behind four fixed occluders. Despite the stimulus remaining constant, observers can alternate between one of four percepts: (1) a continuously morphing cross; (2) two independent perpendicular bars oscillating in depth; (3) a rigidly rotating ellipse observed behind the occluders; (4) a fixed cross observed through a continuously rotating, elliptical aperture. Interestingly, the initial percept naive observers tend to see is percept 1, which is the only nonrigid motion percept. This appears to be a violation of the hypothesized 'rigidity heuristic' in which rigid motion percepts tend to be perceived over retinally equivalent nonrigid ones. Here, we describe the relationships between each of the percepts and the assignment of contour ownership and figure/ground segmentation.

A central function of object recognition is the ability to perceive a moving object as a whole even when it is partially occluded. Because the 2-D image of a partially occluded object is inherently consistent with at least two interpretations (coplanar abutment and occlusion), the visual system must make assumptions about the likely configurations of objects in the world in order to produce a stable percept. In particular, the visual system must take into account whether the observed contour in the image belongs to the occluding surface or the occluded surface.

Movement of the occluded object can both facilitate and hinder recognition. The motion signals generated along continuous portions of contour are inherently ambiguous owing to the aperture problem (Fennema and Thompson 1979; Adelson and Movshon 1982; Marr 1982; Nakayama and Silverman 1988). In contrast, the motions of certain contour features such as line terminators or junctions are potentially unambiguous and can be used to override the aperture problem (Ullman 1979). However, the motions of these trackable features can themselves be misleading if they are due to occlusion and are not intrinsic to the moving contour (Shimojo et al 1989). Thus, in order to generate a stable percept of a moving object, the visual system must determine ownership of any such trackable features, as well as ownership of any moving portion of continuous contour.

In many cases of ambiguous stimuli it has been shown that, of the possible percepts consistent with the retinal image, the one that tends to dominate is the one corresponding to the motion of a rigid object (for example, 'the kinetic depth effect'— Wallach and O'Connell 1953). It has been argued that an assumption of rigid motion is implicitly encoded in the processing of moving stimuli (Ullman 1984). Psychophysical studies of the kinetic depth effect have illustrated the close relationship between contour features (such as junctions, terminators, and regions of high curvature) and the percept of rigid and nonrigid motion, demonstrating that the presence of such trackable features facilitates rigid-motion percepts (Wallach and O'Connell 1953). Studies of anorthoscopic perception have also found a similar relationship between contour features and rigid-motion perception. Anorthoscopic perception is a term used to describe the ability to recognize an object as it moves behind a narrow aperture. The recognition is facilitated by the presence of contour features such as corners or terminators (Zöllner 1862; Helmholtz 1867/1962; Parks 1965). In the absence of such features recognition does not occur, and nonrigid motion is observed within the aperture.

Here, we introduce a new multistable illusion that makes obvious the role of the assignment of contour ownership and determination of implicit and explicit trackable features in the perception of partially occluded moving objects. The illusion also illustrates the relationship between trackable features and rigid-motion percepts. Interestingly, we find that, unlike many ambiguous stimuli in which the rigid percept has been shown to dominate, it is the percept of nonrigid motion that tends to dominate (at least in naive viewers) in the stimulus presented here.

We first encountered the illusion while modifying the well-known 'breathing square' illusion (Meyer and Dougherty 1990, Shiffrar and Pavel 1991), in which a square rotating behind four square occluders is perceived to rotate nonrigidly and undulate in depth. The breathing-square illusion results from an interaction between the motion signals generated by the corners of the rotating square (when exposed) and the motion signal generated by the sides of the rotating square. The interaction between the motion of the square and features defining the occluders has more recently been studied by McDermott and Adelson (2004). They found that form characteristics, such as the concavity or convexity of the occluding contour as well as the presence and orientation of T-junctions, influence both the strength of the occlusion percept and the perception of motion behind the occluder. We were interested to know if an interaction similar to that of the breathing square would occur between the regions of high and low contour curvature located along the contour of an ellipse as it rotated in the 2-D plane behind four occluders (figure 1; also see supplementary video and the color version of figure 2 on the *Perception* website: http://www.perceptionweb.com/misc/p5568/).

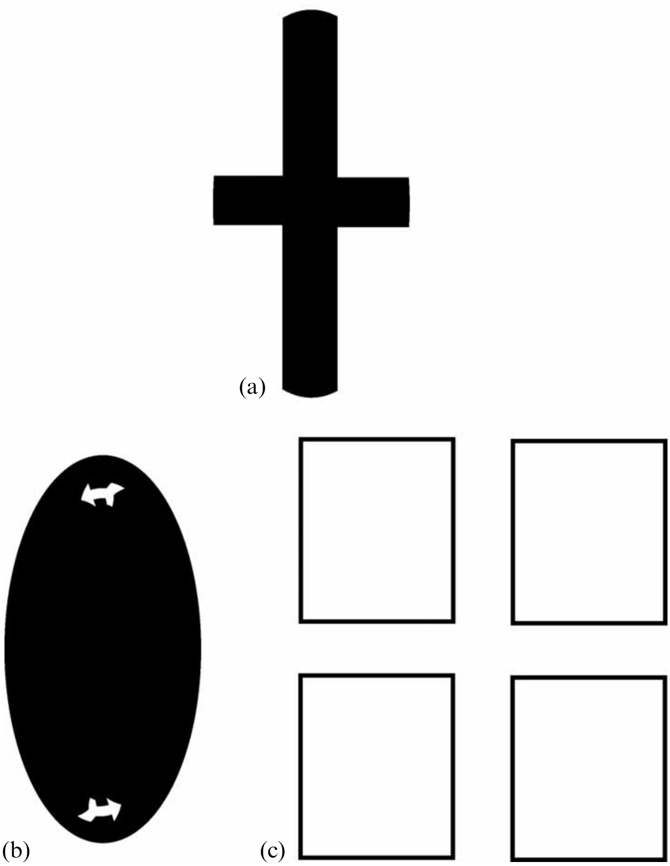

Figure 1. The bar–cross–ellipse stimulus (a) is generated by rotating an ellipse (b) behind four square occluders (c) that are the same color as the background (black outlines not present in actual stimulus).

To our surprise, rather than observing a 'breathing ellipse', we initially observed what looked like a nonrigid cross 'morphing' in the 2-D plane (percept 1: figure 2a). After some time, however, we were clearly able to see an ellipse rigidly rotating behind the four occluders (percept 2: figure 2b). The percept would spontaneously oscillate between these two alternatives with some regularity. Upon further inspection, we observed two more alternative percepts: that of two independent bars oscillating in depth (percept 3: figure 2c), and that of a stationary cross observed through a rotating elliptical aperture (percept 4: figure 2d). We have since presented the stimulus to dozens

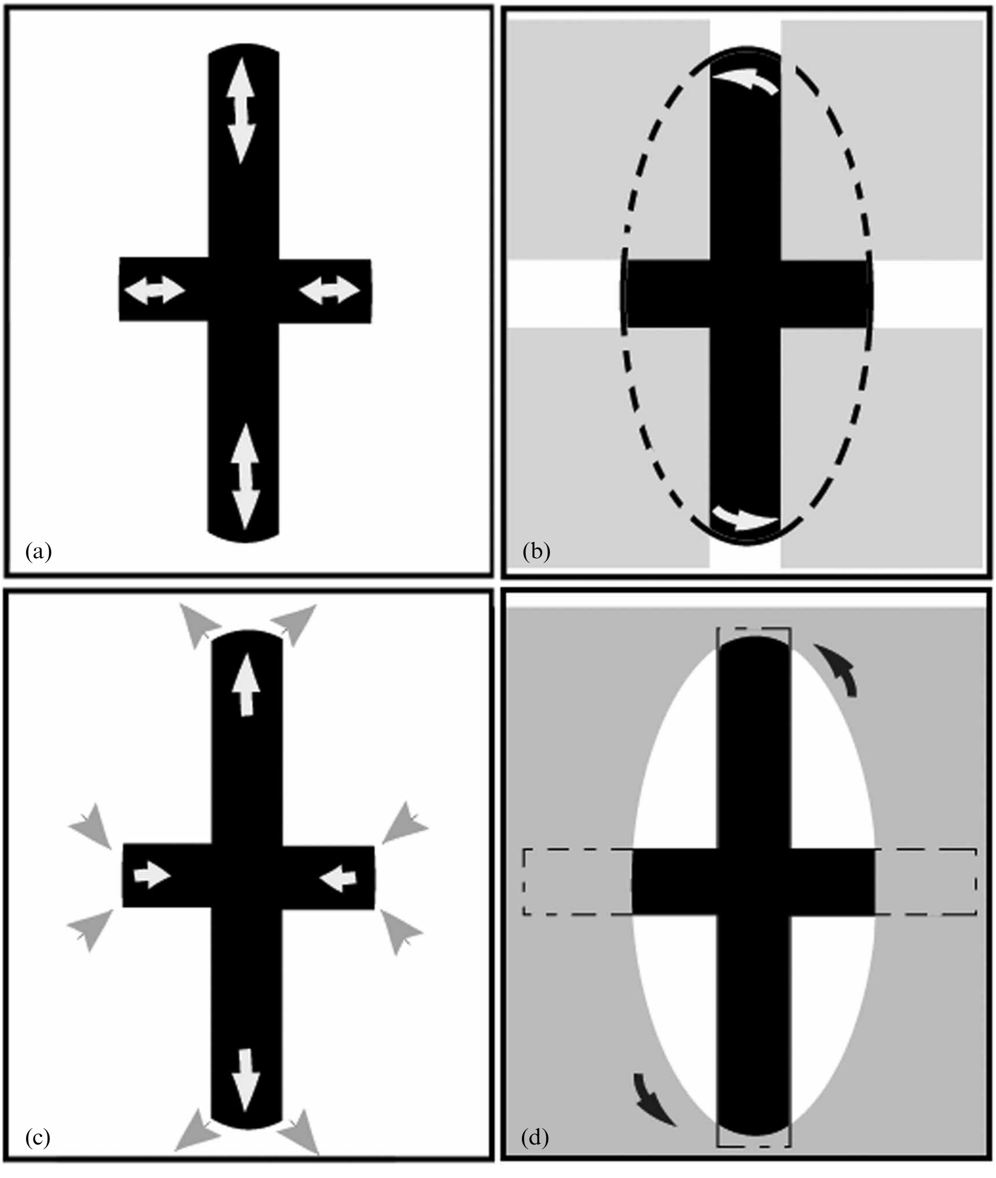

Figure 2. Observers will see one of the following four percepts when viewing the bar – cross – ellipse stimulus: (a) Percept 1: A cross figure nonrigidly morphing in size and shape. (b) Percept 2: An ellipse rigidly rotating behind four square occluders. (c) Percept 3: Two independent perpendicular bars rigidly oscillating in depth. (d) Percept 4: A stationary cross viewed through a rigidly rotating elliptical aperture. A colour version of this figure and a video demonstration can be found on the *Perception* website, at http://www.perceptionweb.com/misc/p5568/.

of undergraduate students. In general, the students will, as we did, initially observe the nonrigid cross. Indeed, in a controlled experiment in which naive viewers were presented with the stimulus for 15 s, thirteen of fourteen reported observing the morphing cross percept. Of these, two reported experiencing a perceptual switch to one of the other percepts (rotating aperture) during the 15 s of presentation. Only one subject did not initially observe the morphing cross (this observer reported seeing the oscillating bars). In general, the alternative percepts are generally observed only after subjects have been instructed to look for them, and, once observed, have a tendency to alternate back and forth between the four possibilities.

Each of the four percepts is consistent with a different figure/ground segmentation, which seems closely related to the assignment of contour ownership and trackable feature motion. In percept 1, the image is segmented into two layers—a figure and a background. All contours in the image are 'owned' by the figure, leading to the percept of an unoccluded cross. The resulting percept of a nonrigid cross is the one that naive observers tend to perceive most dominantly. This would appear to be a violation of the 'rigidity assumption' in which rigid percepts, if possible, tend to dominate. However, the fact that this nonrigid percept tends to dominate is perhaps not so surprising when the role of feature motion is taken into account. Here, the motion percept is primarily driven by the signals generated by the inward/outward motion of the four line-ends (trackable features) rather than those generated along continuous portions of contour. The trackable features in this case are intrinsic to the cross and the resulting nonrigid percept is observed.

In percept 2, the image is segmented into three layers: a frontal layer consisting of the occluders, a middle layer consisting of a partially occluded ellipse, and a background. In this percept, ownership of the contour is divided between the occluders and the ellipse. The 'outer' curved portions of contour are assigned to the occluded figure (ellipse) while the linear 'inner' portions of contour are assigned to the occluding layer (squares). Unlike percept 1, the trackable, inward/outward motion of the line-ends is ignored, while the integrated motion signals of the contour of the ellipse drive the rigid percept of rotation. That this percept is not as readily observed as the morphing cross, speaks to the importance of trackable feature motion in disambiguating motion percepts.

In percept 3, the image is segmented into two figures (perpendicular bars) and a background. The ownership of the contour is divided between the two bars, leading to a percept of one bar partially occluding the other bar. As is the case in percept 1, the trackable feature motion of the line-ends, which get assigned to the bar to which they belong, primarily drives this percept. These motion signals are consistent with the sizes of each bar oscillating out of phase with each other. Here the bars are perceived as being rigidly fixed in their actual size, generating a linear-perspective depth cue that leads to the percept that the bars are, in fact, not changing in size but are, instead, oscillating to-and-fro in depth.

Like percept 2, in percept 4 the image is segmented into three distinct layers: a frontal plane in which an elliptical aperture exists, a middle layer consisting of a partially occluded cross, and a background. In this case, ownership of the contour is again divided between the occluding layer (elliptical aperture) and the figure layer (cross). However, unlike percept 2, the 'inner' linear portions of contour are assigned to the figure (cross) and the 'outer' curved portions of contour are assigned to the occluding layer. As in percept 2, the inward-outward motion of the line-ends is ignored, while the integrated motion signals belonging to the elliptical aperture drive the rigid percept of rotation.

Why is it that, despite the rigid nature of percepts 2, 3, and 4, these percepts do not dominate in naive viewers? Perhaps it is because trackable features, by their very nature, must belong to the object perceived to be moving, and thus have intimately connected with object recognition. It is clear upon inspection of the stationary image depicted in figure 1a that the most readily recognized object is a cross, and therefore the motion of the trackable features is most readily consistent with percept 1. Presumably the strength of each of the four percepts can be manipulated by adjusting the characteristics of the components in the image, thereby strengthening or weakening their likelihood of recognition. For example, percept 2 (rotating black ellipse) would most likely be accentuated by decreasing the sizes of the occluders, thus revealing more of the elliptical contour in the image.

In conclusion, recognition of partially occluded moving objects depends heavily on the assignment of contour ownership and the identification of contour features such as line-ends as being either implicit (belonging to) or explicit (due to occlusion of) the moving object. This multistable illusion exploits these processes by presenting a stimulus whose image is wholly consistent with alternative contour ownership assignments. Here we find, as one might expect, that the dominant percept is one that is primarily driven by trackable features, rather than contour motion signals. Surprisingly though, the resultant percept is one of nonrigid motion. This demonstrates the importance of trackable feature motion as compared to an assumption of rigidity in the disambiguation of motion signals.

Acknowledgments. This project was funded by NIH Grant R03 MH0609660-01 to PUT and by a predoctoral NSF fellowship 2005031192 to GPC.

Gideon P Caplovitz, Peter U Tse
Department of Psychological and Brain Sciences, Moore Hall, Dartmouth College, Hinman Box 6207, Hanover, NH 03755, USA; e-mail: Gideon.Caplovitz@dartmouth.edu, Peter.U.Tse@dartmouth.edu

References

Adelson E H, Movshon J A, 1982 "Phenomenal coherence of moving visual patterns" *Nature* **300** 523 – 525

Fennema C, Thompson W, 1979 "Velocity determination in scenes containing several moving objects" *Computer Graphics and Image Processing* **9** 301 – 305

Helmholtz H von, 1867/1962 *Treatise on Physiological Optics* volume 3 (New York: Dover, 1962); English translation by J P C Southall for the Optical Society of America (1925) from the 3rd German edition of *Handbuch der physiologischen Optik* (first published in 1867, Leipzig: Voss)

McDermott J, Adelson E H, 2004 "The geometry of the occluding contour and its effect on motion interpretation" *Journal of Vision* **4** 944 – 954

Marr D, 1982 *Vision* (New York: Freeman)

Nakayama K, Silverman G H, 1988 "The aperture problem. I. Perception of nonrigidity and motion direction in translating sinusoidal lines" *Vision Research* **28** 739 – 746

Meyer G E, Dougherty T J, 1990 "Ambiguous fluidity and rigidity and diamonds that *ooze*!" *Perception* **19** 491 – 496

Parks T E 1965 "Post-retinal visual storage" *American Journal of Psychology* **78** 145 – 147

Shimojo S, Silverman G H, Nakayama K, 1989 "Occlusion and the solution to the aperture problem for motion" *Vision Research* **29** 619 – 626

Shiffrar M, Pavel M, 1991 "Percepts of rigid motion within and across apertures" *Journal of Experimental Psychology: Human Perception and Performance* **17** 749 – 761

Ullman S, 1979 *The Interpretation of Visual Motion* (Cambridge MA: MIT Press)

Ullman S, 1984 "Maximizing rigidity: the incremental recovery of 3-D structure from rigid and nonrigid motion" *Perception* **13** 255 –2 74

Wallach H, O'Connell D N, 1953 "The kinetic depth effect" *Journal of Experimental Psychology* **45** 205-217

Zöllner F, 1862 "Über eine neue Art anorthoskopischer Zerrbilder" *Annalen der Physik und Chemie* **117** 477 – 484

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 35 2006

www.perceptionweb.com