Chapter 26

Interactions of form and motion in the perception of moving objects

Christopher D. Blair, Peter U. Tse, and Gideon P. Caplovitz

Introduction

This chapter covers a few highlights from the past 20 years of research demonstrating that there is 'motion from form' processing. It has long been known that the visual system can construct 'form from motion.' For example, appropriate dot motions on a two-dimensional computer screen can lead to a percept of, say, a rotating three-dimensional cylinder or sphere. Less appreciated has been the degree to which perceived motion follows from processes that rely upon rapid analyses of form cues. Percepts that depend on such form-motion interactions reveal that form information can be processed and integrated with motion information to determine both the perceived velocity and shape of a moving object. These integration processes must be rapid enough to occur in the brief period, probably less than a quarter of a second, between retinal activation and visual experience.

Data suggest that global form analyses subserve motion processing in at least five ways (Porter et al., 2011). Here, we describe three examples in which the analysis of form significantly influences our experience of moving objects. The following examples have been chosen not only for their distinctiveness, but also to compliment other examples described in detail within other chapters of this book (Bruno & Bertamini, 2013; Herzog & Öğmen, 2013; Hock, 2013; Vezzani et al., 2013). First, we describe Transformational Apparent Motion, a phenomenon that reveals how form analyses permit the figural segmentation dedicated to solving the problem of figure-to-figure matching over time (Hsieh and Tse, 2006; Tse, 2006; Tse & Caplovitz, 2006; Tse & Logothetis, 2002). Secondly, we describe how the size and shape of an object can influence how fast it is perceived to rotate. These interactions reveal the way in which form analyses permit the definition of trackable features whose unambiguous motion signals can be generalized to ambiguously moving portions of an object to solve the aperture problem (Caplovitz et al., 2006; Caplovitz & Tse, 2007a,b). Finally, we describe a number of peculiar ways in which the motions of individual elements can interact with the perceived shape and motion of a global object constructed by the grouping of these elements. These phenomena reveal that the form analyses that underlie various types of perceptual grouping can lead to the generation of emergent motion signals belonging to the perceptually grouped object that appear to underlie the conscious experience of motion (Caplovitz & Tse, 2006, 2007b; Hsieh & Tse, 2007; Kohler et al., 2010; Kohler et al., 2009).

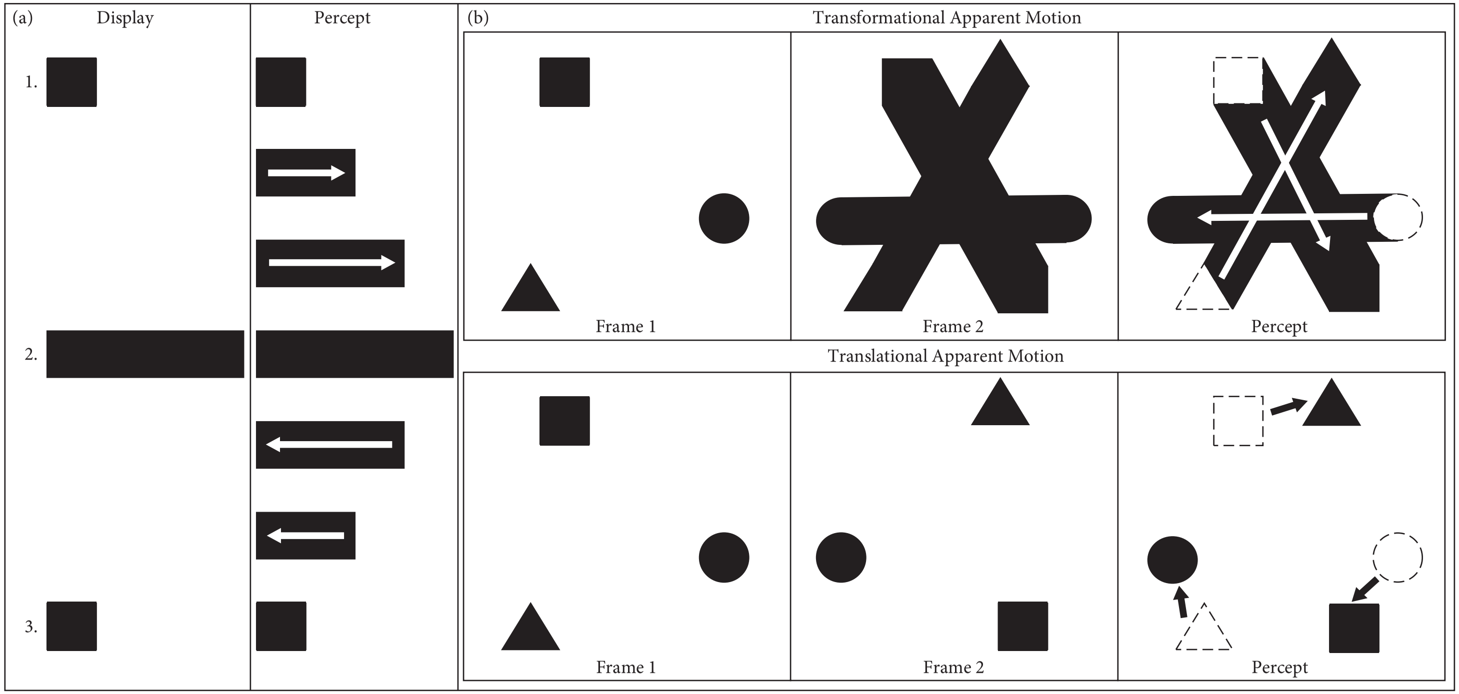

Fig. 26.1 (A) Transformational Apparent Motion (TAM). Two abutting shapes are flashed in sequence, as shown on the left. The resulting percept is of one shape smoothly extending from, and retracting back into the other, as depicted on the right. (B) TAM v. Translational Apparent Motion. In TAM displays (top), when two frames are flashed in sequence, if the shapes in the second frame abut those in the first frame the percept is of smooth deformation that is based on the figural parsing of the objects in both frames. However, in translational apparent motion displays (bottom), when the shapes in the second frame do not abut those in the first frame, rigid motion to the nearest neighbor is perceived independent of any figural parsing.

OUP UNCORRECTED PROOF – FIRSTPROOFS, Mon Nov 03 2014, NEWGEN

Transformational Apparent Motion

Background

A phenomenon known as Transformational Apparent Motion (TAM) has received much attention over the past 20 years and sparked a renewed examination of the role of form analyses in high-level motion processing. TAM occurs when two shapes, overlapping in space, are presented at different points in time, giving the illusion that one shape smoothly transforms into the other (Tse et al., 1998). Precursors to TAM included 'polarized gamma motion' and 'illusory line motion,' with the latter being a rediscovery and re-examination of the first (Hikosaka et al., 1991, 1993a,b; Kanizsa, 1951, 1979). A classical demonstration of polarized gamma motion and illusory line motion is illustrated in Figure 26.1A. Illusory line motion arises when a horizontal bar is presented shortly after a transient cue located at one end of the bar. When this occurs, the bar appears to extend out from the cue, rather than appearing all at once. Thus, rather than the sudden appearance of a stationary object, a motion percept is observed in which an object appears to morph from one shape to another.

An initial hypothesis for why these phenomena occur posited a primary role for attention. Specifically, the sudden onset of the cue stimulus possibly draws attention and establishes an attentional gradient that extends outward from the cue location. Because information at attended locations was presumed to be processed faster than at unattended locations, the target stimulus would be processed asynchronously, leading locations closer to the center of the attentional gradient to reach conscious awareness prior to those located more distally. This would thereby lead to the illusory percept that the horizontal bar successively extends out from the point of attention (Faubert and von Grünau, 1995; Stelmach and Herdman, 1991, Stelmach et al., 1994; Sternberg and Knoll, 1973; Titchener, 1908; von Grünau and Faubert, 1994). While attentional gradients may, indeed, play some role in the illusory percept, subsequent experimentation suggested a dominant contribution of other factors. For example, TAM can be observed even when attention is allocated away from the cue. Also, if two cues – a red and a green dot – are presented simultaneously, some distance apart, when a red line appears abutting each cue and between them, the line always appears to extend from the red dot, regardless of which cue is originally attended (Downing and Treisman, 1995, 1997; Hsieh et al., 2005; Tse and Cavanagh, 1995; Tse et al., 1996, 1998).

To account for these non-attentional effects, it has been argued that the illusory motion observed in these stimuli arises from figural parsing (Tse et al., 1998; Tse and Logothetis, 2002). Figural parsing occurs when contour and surface relationships are compared across successive scenes. Thus, based on their relative surface and contour relationships, the visual system determines which shapes viewed at one time point correspond to which shapes viewed at a subsequent time point. In the case of TAM, the visual system infers that an existing figure has changed its shape into that of the new figure, leading to the perception of continuous deformation. Implicit in this hypothesis is a fundamental role for form processes that extract information about the shape and surface characteristics of objects. Moreover, as the motion percept in TAM displays depends upon the output of these processes, this processing must occur either prior to, or coincident with motion processing. In this view, processes that represent form information help solve the 'what went where?' question of object movement. This occurs in two steps. First, individual objects are identified or 'parsed' in a scene. The second step involves matching these parsed objects to the objects present in the preceding scene.

The processes underlying TAM can be contrasted to those underlying classical translational apparent motion. In classical translational apparent motion, when there are multiple objects in both the first and second scene, motion correspondences tend to be formed between spatially-proximal objects. This is true even if the proximal objects have dramatically dissimilar shape and surface characteristics. As with TAM, this would imply that the object had grossly deformed from one scene to the next. However, this deformation is determined not on the basis of object parsing and figural matching, but rather on the basis of spatiotemporal proximity (Ullman, 1979). As such, observations such as these led to the discounting of the importance of form features in determining object motion in the past (Baro and Levinson, 1988; Burt and Sperling, 1981; Cavanagh and Mather, 1989; Dawson, 1991; Kolers and Pomerantz, 1971; Kolers and von Grünau, 1976; Navon, 1976; Ramachandran et al., 1983; Victor and Conte, 1990). However, as illustrated in Figure 26.1B, TAM can still be observed in cases where the nearest neighbor principle may be violated in favor of matching shapes across scenes that actually comprise more distant figures. This has been demonstrated to result from a set of parsing and matching principles involving the analysis of contour relationships among successive and abutting figures (Tse et al., 1998; Tse and Logothetis, 2002). This appears to result largely from an analysis of good contour continuity, which indicates maintained figural identity, and contour discontinuity, which implies figural differences. Given the lack of figural overlap in most translational apparent motion displays, this parsing is generally unnecessary in determining 'what went where?'

Neural correlates

Functional magnetic resonance imaging has determined which areas of the brain show the greatest blood oxygen level dependent (BOLD) activity in response to TAM displays, as compared with control stimuli (Tse, 2006). Using a region of interest analysis, this study found greater activity in response to TAM than control displays in V1, V2, V3, V4, V3A/B, hMT+, and the Lateral Occipital Complex (LOC). An additional whole-brain analysis identified an area in the posterior fusiform gyrus that was also found to be more active during the perception of TAM than control stimuli. The recruitment of early retinotopically organized areas highlights the importance of the basic visual processes (i.e. spatially specific detection of edges and contour features) that underlie the perception of TAM. The recruitment of higher-level areas likely reflects the more global processing that must underlie figural parsing and subsequent figural matching.

Of particular interest is the recruitment of the LOC. The LOC is now fully established as playing a fundamental role in form processing and object recognition (Grill-Spector et al., 2001; Haxby et al., 2001; Kanwisher et al., 1996; Malach et al., 1995) and, like TAM, has been shown to process global 3D object shape, as opposed to just local 2D shape features (Avidan et al., 2002; Gilaie-Dotan et al., 2001; Grill-Spector et al., 1998, 1999; Malach et al., 1995 Mendola et al., 1999; Moore and Engel, 2001; Tse and Logothetis, 2002; Kourtzi and Kanwisher, 2000, 2001; Kourtzi et al., 2003a). A reasonable interpretation of the increased activity in LOC during the viewing of TAM displays relative to control stimuli is that in addition to processing global form and figural relationships, this information is also output to motion-processing areas of the brain, such as hMT+.

Given this interpretation, and the increased activity demonstrated in both LOC and hMT+ during TAM displays, it seems that hMT+ and LOC, rather than being motion processing and form processing areas, respectively, may both serve as part of a form/motion processing circuit. In fact, multiple studies have shown functional and anatomical overlap between LOC and hMT+ (Ferber et al., 2003; Kourtzi et al., 2003a; Liu and Cooper, 2003; Liu et al., 2004; Murray et al., 2003; Stone, 1999; Zhuo et al., 2003). As noted later in this chapter, it is likely that V3A/B, an area that also shows increased activity in response to TAM displays, plays a key role in this form/motion processing circuit. These findings call into question the traditional view of separate motion and form processing streams contained in the dorsal 'where' and ventral 'what' pathways (Goodale and Milner, 1992; Ungerleider and Mishkin, 1982). Although at the very highest representational levels 'what' and 'where' may be largely independent (Goodale and Milner, 1992; Ungerleider and Mishkin, 1982), form and motion processes are likely to be non-independent within the processing stages that serve as inputs to these later representations.

Additional work has been done using electroencephalography (EEG) to study visually-evoked potentials (VEP) in response to TAM displays as compared with displays that only flashed, but lacked the TAM percept (Mirabella & Norcia, 2008). This study found that the VEP waveform evoked by pattern onset and offset was significantly more symmetrical for TAM displays than for flashing displays. The timing of such TAM-related processing appears within the first 150 ms of object appearance and disappearance, once again implicating the involvement of early visual areas in processing TAM. Furthermore, it was shown in the frequency domain that there was a noticeable reduction in the odd-harmonic components in the frequency spectra for the TAM display, as compared with that for a flashing patch alone. This further reflects the increased symmetry in the TAM VEP waveform. Interestingly, as the contrast between the cue and flashing patch in the TAM display was increased, the symmetry in the resulting VEP waveform decreased. Behavioral data matched this observation, as the likelihood of participants perceiving TAM in the display was strongly correlated with the symmetry of the VEP waveform. Thus, both behavioral and EEG data further demonstrate the influence of object surface features on perceived movement.

Implications for Models of Transformational Apparent Motion

The only formal model that we are aware of that attempts to account for TAM involves three interacting subprocesses (Baloch and Grossberg, 1997). The first is a boundary completion process where activity flows from V1 to interstripe V2 to V4. The second is a surface filling process where activity flows from blob V1 to thin stripe V2 to V4. The third is a long-range apparent motion process where activity flows from V1 to MT to MST. The model includes an additional link between V2 and MT that allows the motion-processing stream to track emerging contours and filled-in color surfaces (Baloch and Grossberg, 1997). The model represents a locally-based, bottom-up explanation of TAM. In the fMRI experiment described above, each of the areas referenced in the model has shown higher relative activity during the viewing of TAM displays. However, the model fails to account for increased activity shown in V3v, V3A/B, and LOC. Furthermore, TAM has been shown to be influenced by global configural relationships among stimuli, which this locally based model cannot explain (Tse and Logothetis, 2002). TAM demonstrates many of the central problems that the visual system must solve, which have been the subject of much study in the field of visual neuroscience: How is local form information integrated into a global representation of spatiotemporal figural relationships, and how does this, in turn, influence the interpretation of local features (Kenkel, 1913; Wertheimer, 1912/1961)? During the perception of TAM, figural contours must be analysed and integrated globally, over both space and time within and between scenes.

For both contour integration in general and TAM, fMRI studies have demonstrated the strongest activity in lateral occipital areas of both the human and monkey brain (Altmann et al., 2003; Kourtzi et al., 2003b; Tse, 2006). However, both V1 and V2 also show increased activity during such processes (Altmann et al., 2003; Caplovitz et al. 2008; Kourtzi et al., 2003b; Tse, 2006). While increased activity in V2 may be unsurprising, given that single unit recordings have shown its involvement in the perception of illusory contours (von der Heydt et al., 1984), no such involvement as early as V1 had previously been demonstrated. In more recent years, visual areas V1 and V2 have been implicated in the processing of global shape (Allman et al., 1985; Fitzpatrick, 2000; Gilbert, 1992, 1998; Lamme et al., 1998) despite the traditional view that V1 is only involved in the processing of local features (Hubel and Wiesel, 1968). However, it is still unclear whether such activity in V1 results from bottom-up or top-down activation. A recent fMRI study found increased activity in response to the spatial integration of individual elements into perceptually grouped wholes in early visual cortex, possibly as early as V1 (Caplovitz et al., 2008). This was true, despite each individual element being located in the periphery of a different visual quadrant, suggesting such increases in activity are likely due to top-down feedback.

Separate from TAM, parsing can be important in other standard and apparent motion displays, as pooling the motion energy of multiple objects moving through the same point in space would lead to inaccurate motion signals (Born and Bradley, 2005). Motion signals arising at occlusion boundaries may also be spurious (Nakayama and Silverman, 1988), and parsing can facilitate the segmentation of spurious from real motion signals. It would appear that the visual system possesses such parsing mechanisms and they help us to accurately perceive the motion of multiple overlapping objects (Hildreth et al., 1995; Nowlan and Sejnowski, 1995). While there is evidence that hMT+ plays some role in such motion parsing processes (Bradley et al., 1995; Stoner and Albright, 1992, 1996), other evidence suggests that aspects of this process, such as figure segmentation, do not take place in hMT+. Rather, it is more likely that specialized areas, such as LOC handle global figural segmentation and similar processes, and that the resulting neural activity is then output to hMT+. Given such an interaction, the analyses of form and motion, and thus shape over time and space, can be seen as interacting inseparable processes. That form and motion should be analyzed in an integrated spatiotemporal fashion was suggested as early as 1979 (Gibson), and has been re-emphasized in more recent years (Gepshtein and Kubovy, 2000; Wallis and Bülthoff, 2001).

Size, Shape and the Perceived Speed of Rotating Objects: Trackable Features

Recent research has demonstrated that the shape of an object directly affects the speed with which it appears to rotate (Blair, Goold, Killebrew & Caplovitz, 2013; Caplovitz et al., 2006; Caplovitz and Tse, 2007a; Porter et al., 2011). Specifically, objects with distinctive contour features, such as corners or regions of high or discontinuous contour curvature are perceived to rotate faster than those without such contour features. For example, when ellipses of various aspect ratios are rotated with the same angular velocity, the 'skinnier' an ellipse is, the faster it appears to rotate (Caplovitz et al., 2006).

There are various explanations for why this may be the case, and experiments have been conducted to dissociate between them. For example, skinnier objects in general may appear to rotate faster than fatter ones. Such an explanation is rooted in the temporal frequency with which contrast changes at any particular location in the visual field, highlighting the intrinsic ambiguity that arises between spatial frequency, speed, and temporal frequency (Brown, 1931). Simply put, the surface of a rotating skinny object will sweep across a neuron's receptive field in less time than that of a fatter object. This hypothesis can be ruled out by the fact that no differences in perceived speed were observed between the perceived speed of skinny and fat rectangles (Caplovitz et al., 2006).

A second hypothesis is that distinctive contour features serve as trackable features that provide an unambiguous source of information about the speed and direction of motion of a given object.

This hypothesis is rooted in the works of Wallach (Wallach, 1935; Wallach & O'Connell, 1953; Wallach et al., 1956) and Ullman (1979), which highlight the importance of such form features in extracting 3D structure from motion (i.e. the Kinetic Depth Effect). In the case of a skinny ellipse, the regions of high curvature located at the ends of the major axis may serve as an additional source of motion information that is unavailable in the case of a fat ellipse. Moreover, this hypothesis is consistent with the lack of effect observed with rotating rectangles whose corners may act as trackable features regardless of whether they belong to a skinny or fat rectangle. To directly test this hypothesis, an experiment was conducted in which the corners of a rectangle were 'rounded off' to a lesser or greater degree (Caplovitz et al., 2006). The more the corners were rounded, the slower the rounded-rectangle appeared to rotate, thereby providing strong support in favor of the form-defined trackable features hypothesis (see Figure 26.2A).

A third hypothesis, and one consistent with the data derived from the experiments described above, is that the perceived speed of a rotating object is determined by the magnitudes of locally detected 1D motion signals (Weiss and Adelson, 2000). Changes to an object's shape will change the distribution of component motion signals detected along its contour. When the magnitudes of component motion signals derived from a skinny ellipse were compared with those derived from a fat ellipse (see Figure 26.2B) it was found that they scaled in a manner wholly consistent with the changes in perceived speed. Moreover, because the magnitudes of component motion signals scale

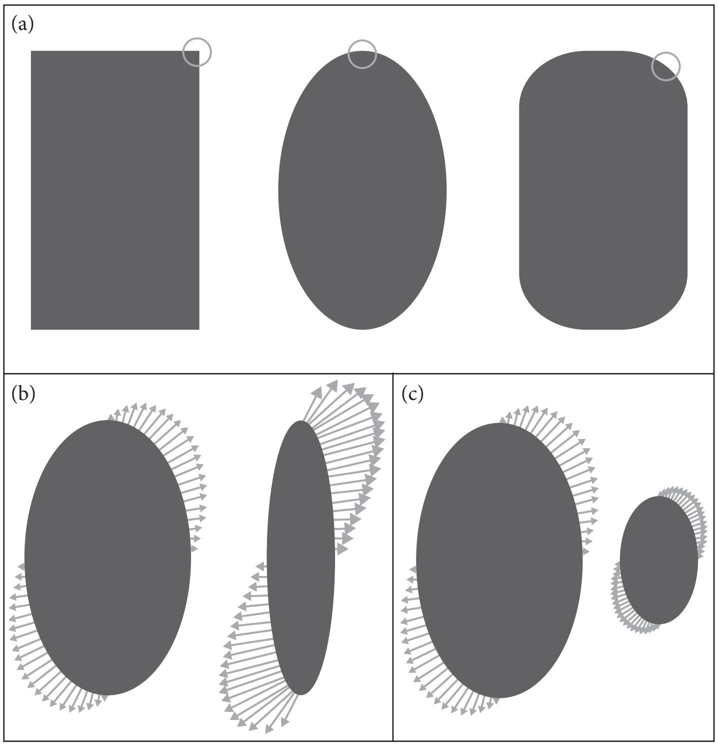

Fig. 26.2 Trackable features and component vectors. (A) Proposed trackable features on rectangles, ellipses, and rounded rectangles. (B) Changes in local component motion vectors of a rotating ellipse as a function of changes in aspect ratio. (C) Changes in local component motion vectors as a function of changes in the size of rotating objects.

as a function of their distance from the center of rotation, there are no differences in distribution of such signals between skinny and fat rectangles. Although the relationship between component motion magnitude and perceived speed is not as precise for the case of the rounded rectangles, there is indeed a parametric decrease in the local distribution of component motions signals in the corner regions as the corners become more and more rounded (Caplovitz et al., 2006).

As such, these initial sets of experiments were unable to conclusively determine whether shape-related changes in perceived rotational speed arise due to trackable features or the integration of local component motion signals. It was not until very recently that experiments were conducted to explicitly dissociate between these two hypotheses (Blair et al., 2013). This study specifically examined the case of angular velocity. A hallmark of angular velocity is that it is size invariant. Making a rotating object smaller will not change its angular velocity. However, doing so will systematically decrease the magnitudes of the component motion signals derived along its contour (see Figure 26.2C). The study compared the perceived rotational speeds of small and large objects. There were two primary findings of the study: first, across a range of object categories: ellipses, rectangles, stars, and rounded rectangles, smaller objects appear to rotate more slowly than larger objects. This finding is what would be predicted by the local-motion integration hypothesis. However, the second main finding of the study is that the degree to which smaller objects appear to rotate slower is dependent upon the shape of the object. Specifically, while the relative change in perceived speed of rectangles with very rounded corners is nearly perfectly predicted by the relative magnitudes of the component motion signals, very little change in perceived speed is observed for regular rectangles, skinny ellipse, and star-shapes. Indeed, simply reducing the degree to which the corners of the rounded-rectangles were rounded off reduced the effect size of perceived rotational speed. These two findings suggest that both hypotheses are likely to be true: the perceived speed of a rotating object is determined by a combination of locally detected motion signals, which comprise a scale-variant source of information, and the motion of form-defined trackable features, which comprise a scale-invariant source of information.

What is important to note is that both sources of information are shape-dependent. However, only the trackable feature motion requires an analysis of form, because in order to provide a useful source of information, the trackable feature must first be classified as belonging to the object that is rotating (see figure parsing above). Moreover, the motion of the trackable feature must be attributed to other locations along the object's contour. Lastly, in order to produce a size-invariant representation (i.e. angular velocity), the motion of a trackable feature must be integrated with information about its distance from the center of rotation, a necessarily non-local computation. In the case of objects that simultaneously translate as they rotate, it appears to be the case that the rotational motion around the object's center is segmented from the overall translational motion of the object (Porter et al., 2011). This suggests that the size invariant signal derived from the motion of a trackable feature involves the computation of the object's center.

The effects of object shape on the perceived speed of rotational motion have also been observed and examined in the context of motion fading. Motion fading occurs when a slowly drifting or rotating pattern appears to slow down and then momentarily stop, while the form of the pattern is still visible (Campbell and Maffei, 1979, 1981; Lichtenstein, 1963; Spillman and De Weerd, 2003). Experiments have shown that the presence of trackable features extends the time that it takes motion fading to occur for rotating objects, as compared with those rotating objects, which do not possess distinct trackable features (Hsieh and Tse, 2007). Furthermore, if the trackable features of objects such as ellipses are made even more distinct by increasing a rotating ellipse's aspect ratio, it takes even longer for motion fading to occur (Kohler et al., 2010). It was further shown that the effect of shape on the time for motion fading to occur is mediated by the perceived speed of the rotating object. For example, a fatter ellipse will appear slower than a skinny ellipse and will therefore take less time for motion fading to occur. Thus, by influencing the perceived speed of rotation, an object's contour features dictate how long it takes for a slowly rotating object to appear to cease moving. This demonstrates the importance of the form-motion interaction that underlies the role of trackable features role in the perception of rotational motion. Not only do they provide a direct effect on perceived speed, but also indirect effects on other aspects of motion perception.

Neural correlates

Clearly, there is strong behavioral evidence for the existence of multiple form–motion interactions. The question stands: where in the brain might these interactions take place? In the context of the role form plays in the perceived speed of rotating objects, evidence from fMRI studies has implicated the involvement of V3A. When shown rotating objects that modulated their contour curvature at one point while remaining constant in speed and area, BOLD activity was also modulated in area V3A of observers' brains (Caplovitz & Tse, 2007b). Previous research focused on this area has led to findings consistent with the interpretation that V3A makes use of areas of contour curvature to process the rotational motion of objects. For one, it has been shown in several studies that area V3A is motion selective (Tootell et al., 1997; Vanduffel et al., 2002). Motion processing is only half of the story, and sure enough, V3A per cent BOLD signal change has also been correlated with contour and figural processing, even when contours and figures are not consciously perceived (Schira et al., 2004). To go a step further, BOLD activity in V3A has been correlated with various additional form-motion interactions. Specifically, it has been shown multiple times that there is a greater percent BOLD signal change in the V3A when participants observe coherent, as opposed to random motion (Braddick et al., 2000, 2001; Moutoussis et al., 2005; Vaina et al., 2003). Finally, it was found that the V3A is more responsive to rotational than translational motion (Koyama et al., 2005). In combination, these various findings indicate that V3A makes use of form information, specifically contour curvature, to process motion information about moving objects. The strongest activity may result in situations where the motion is more difficult for the visual system to interpret, such as with rotation (Kaiser, 1990).

Neurophysiological data recorded in area MT of macaques has further elucidated some specifics of how areas of contour curvature on objects may be used in processing object motion. Specifically, certain neurons in macaque MT have been shown to respond more to the terminator motion of lines than to the ambiguous motion signals present along a line's contour. In addition, these neurons respond strongest when terminators are intrinsically owned, as opposed to when they are extrinsic (Pack et al., 2004). Interestingly, this process is not instantaneous, as it takes roughly 60 ms for neurons in macaque MT to shift their response properties from those consistent with motion perpendicular to a moving line, regardless of its actual direction of motion, to those consistent with the true motion of the line independent of its orientation (Pack and Born, 2001). Behavioral data examining initial pursuit eye movements support this finding, in that observers will initially follow the motion perpendicular to the moving line before then exhibiting eye movements that follow the unambiguous motion of line terminators. Further neurophysiological evidence has indicated that neurons of this sort (dubbed end-stopped neurons) may be present in the visual system as early as area V1 (Pack et al., 2003). This would mean that trackable feature information could be extracted and utilized as early on as V1 in the visual processing stream. All these findings could help explain how the visual system is capable of overcoming the aperture problem under various circumstances using trackable features, and also, why it does not always do so perfectly.

From Moving Parts to Moving Wholes: the Perceived Motion of Perceptually Grouped Objects

Just as an object's shape has been shown to affect its perceived motion, additional processes, such as perceptual grouping and the formation of contours from discrete elements, can lead to changes in perceived motion. For example, one study examined how the perception of the speed for rotating ellipses was modulated when the ellipses' contours were constructed using individual dots, instead of a continuous contour (Caplovitz & Tse, 2007a). Under these circumstances, one might expect that changing the aspect ratios of these ellipses should have no effect on their perceived speed, as the individual dots should serve as unambiguous trackable features not subject to the aperture problem. However, this was only the case if the dots were spaced sufficiently far apart. While not in direct contact with one another, when spaced closely enough together, aspect ratio-related changes in perceived speed were observed. This was true even when the ellipses were formed using contrast-balanced dots that minimally activate neurons sensitive to low-spatial frequencies to whose large receptive fields closely spaced dots may produce similar patterns of activity as a continuous contour. It was subsequently hypothesized that when the dots are closely spaced the visual system is incapable of following the motion of a single dot. In the absence of such locally unambiguous motion, the visual system makes use of the information from the perceptually grouped contour implicit in the dot arrangement (Caplovitz & Tse, 2007a).

Further evidence for the effects of grouping on perceived motion has been demonstrated using the motion fading paradigm. Specifically, when elements are part of a slowly rotating display, if disparate elements can be grouped in such a way as to form the perception of an object that possesses trackable features, the amount of time necessary for motion fading to occur is increased (Hsieh & Tse, 2007; Kohler et al., 2010). Similar to the previously described experiment examining the perceived rotational speed of dotted ellipses, the aspect ratio of such ellipses affects the time course of motion fading only when the dots are spaced closely enough that a single dot cannot be tracked by the visual system (Kohler et al., 2010).

While these previously discussed examples of the effects of grouping on motion perception appear to be largely automatic in nature, multistable perceptions involving grouping and perceived speed have also been demonstrated. Specifically, if four dot pairs are evenly spaced in a square formation, and each pair rotates around a common center, observers may interpret the movement as four rotating dot pairs, or two flat squares moving in a circular motion with one in front and the other behind, the dots in the pairs making up their corners (Anstis, 2003; Anstis and Kim, 2011). As a participant's perception and interpretation changes, so does the perceived speed of elements present (Figure 26.3A). When perceptually grouped into the global percept of a square, the perceived speed of the display appears to slow down (Kohler, Caplovitz, & Tse, 2009). The dots may be exchanged for various elements that bias the perception in one direction or another (Figure 26.3B). Such elements have been shown to be perceived as moving faster when viewed simply as rotating pairs, than when seen as being part of any of the illusory shapes that may result from interpreting them as being corners instead of individual elements (Kohler, Caplovitz, & Tse, 2009). Thus, form information resulting from both automatic and multistable perceived groupings of moving objects can affect the perceived motion of such groups.

Thus far, the effect of object shape on its perceived motion has been principally discussed. However, there are also examples showing that the movement of an object can influence its perceived shape (i.e. the Gelatinous Ellipse, Weiss and Adelson, 2000; and the Kinetic Depth Effect,

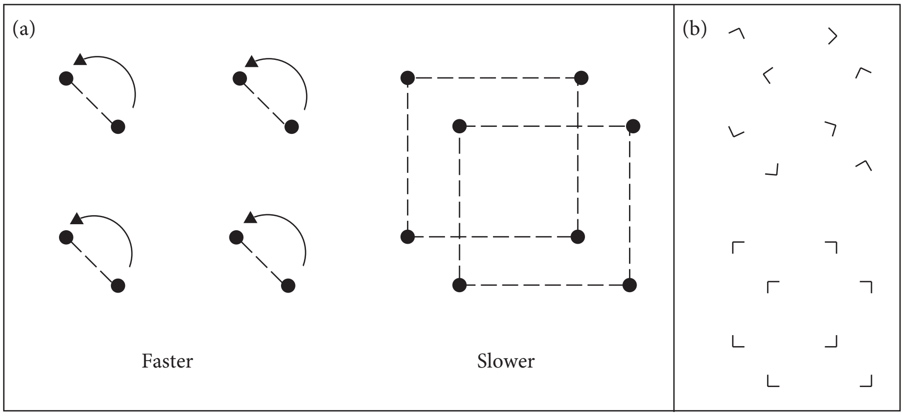

Fig. 26.3 Emergent motion on the basis of perceptual grouping. (A) When four dot pairs, each pair rotating around its own common center, are perceived as separate objects, they are perceived to rotate faster than when dots are perceived to form the corners of two squares translating in a circular pattern with one sliding in front of the other. (B) The percept of individual elements or square corners may be biased by element shape and arrangement, with individual elements most likely to be seen when misaligned (top), and squares more likely to be seen when the elements are aligned (bottom).

Wallach and O'Connell, 1953). Recently, it has been demonstrated that the movement-dependent shape distortions can come as a result of local form/motion interactions in elements grouped to form a larger perceived object. As previously mentioned, elongated objects are perceived to move faster when moving in a direction parallel, as opposed to orthogonal, to their elongated axis (Georges et al., 2002; Seriès et al., 2002). Taking advantage of this observation, an experiment was conducted in which differentially elongated Gaussian blobs were used to form the corners of illusory four-sided translating shapes. In the experiment, the blob would be orientated such that those on the leading edge of the illusory object would be either parallel or orthogonal to the direction of motion and those on the trailing edge of the illusory shape would be orientated orthogonally to those on the leading edge. It was found that when those on the leading edge were parallel to the direction of motion, the resulting illusory object appeared to be elongated, while the opposite effect was observed when blobs on the leading edge were oriented orthogonally to the direction of motion, as depicted in Figure 26.4 (McCarthy et al., 2012). This example reveals how form and motion interact with each other across a range of visual processing stages from very early (local orientation dependent perceived speed) to later representations of perceived global shape.

As mentioned in the introduction, a 3D representation of a moving object can be derived from appropriate 2D velocities of seemingly random dot displays. In such form-from-motion displays, depth, 3D object shape, and 3D object motion may be perceived if seemingly random dot fields are moved in ways consistent with the dots in motion being affixed to a particular 3D shape (Green, 1961). This process represents a form of perceptual grouping in which the individual dots are grouped into a single perceptual whole. Intriguingly, the shape and motion of the perceived object do not always match what would be predicted based upon the individual motions of the dots that make up the display. Instead, characteristics of the shape and motion of the global object depend upon the shape and motion of the object itself. For example, perceived variations in the angular velocity of rotating 3D shapes simulated by dot fields were more closely tied to the

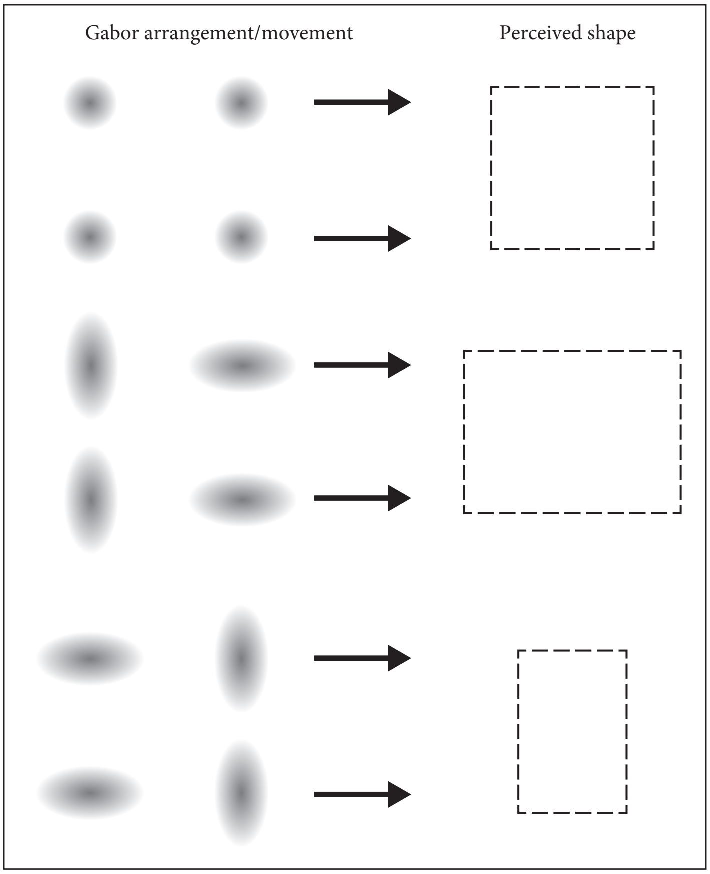

Fig. 26.4 Form-motion-form interaction. When elliptical Gaussians are arranged in a square formation and translated in a single common direction, if the leading edge and trailing edge Gaussians are orientated 90° from one another, the perceived moving shape will appear to be a rectangle instead of a square. The shape will appear elongated if the leading edge Gaussians are orientated parallel to their direction of translation, and compressed if the leading edge Gaussians are orientated orthogonal to their direction of translation.

perceived deformation of the rotating shapes than on actual variations in their angular velocities (Domini et al., 1998). Similarly, the perceived slant of a simulated surface varies as a function of the angular velocity with which it rotates when other factors are kept constant (Domini & Caudek, 1999). These various effects have been demonstrated both when objects are rotated, while being passively observed, and when object motion is a function of simulated optic flow in response to observer movement (Caudek et al., 2011; Fantoni et al., 2010, 2012). Additionally, even when binocular visual cues such as disparity are available, such biases and misperceptions are still observed (Domini et al., 2006). The perception of these effects and visual biases is also correlated with changes in grasping movements for the simulated objects (Foster et al., 2011). A model based on the assumption that the analysis of 3D shape is performed locally accounts well for successful and unsuccessful interpretation of 3D shape and the movement of 3D shapes by human observers, as demonstrated by a variety of form motion interactions observed using this paradigm (Domini & Caudek, 2003). Thus, not only is visual perception affected by form motion interactions, but the practical behaviors in response to such perceptions are also adjusted accordingly.

Conclusion

These results can be taken as further evidence for the inherently constructive nature of motion processing, and the importance of form operators in motion processing. While it is not clear where in the brain the analysis of form occurs that results in the perception of rotational motion, it probably occurs within some or all of the neural circuitry that realizes the form–motion interactions described above. These results support the general thesis that there are, broadly speaking, two stages to motion perception – one, where motion energy is detected by cells in early visual areas tuned to motion magnitude and direction, and another stage where this detected information is operated upon by grouping and other visual operators that then construct the motion that will be perceived (Caplovitz & Tse, 2007a; Hsieh & Tse, 2007; Kohler et al., 2009, 2010). This means that perceived motion, while constructed on the basis of locally detected motion information, is not itself detected or even present in the stimulus. It should also be noted that, while we have focused on specific examples from only three broad categories of form motion interaction, these examples represent only a small subset of what has been identified and tested at this time with further examples ranging as far as the processes underlying the perception of biological motion and how motion is conveyed through static images (i.e. motion streaks).

Classically, form and motion perception were considered to be mediated by independent processes in the visual system. Indeed there is a good deal of evidence for such independence at the earliest stages of visual processing, as well as at the highest levels of perceptual representation. However, there is growing evidence suggesting that the mechanisms that process form and motion characteristics of the visual scene mutually interact in numerous and complex ways across a range of mid-level visual processing stages. These form-motion interactions appear to help resolve fundamental ambiguities that arise at the earliest stages in the processing of the retinal image. By combining information from both domains, these form motion interactions allow potentially independent high-level representations of an object's shape and motion to more accurately reflect what is actually occurring in the world around us.

Acknowledgment

This work was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number 1P20GM103650-01, and a grant from the National Eye Institute: 1R15EY022775.

References

Allman, J. M., Miezin, F., and McCuiness, E. (1985). Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Ann Rev Neurosci* 8: 407–430.

Altmann, C. F., Bülthoff, H. H. and Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Curr Biol* 13(4): 342–349.

Anstis, S. (2003). Levels of motion perception. In *Levels of Perception*, edited by L. Harris & M. Jenkin, pp. 75–99. New York: Springer.

Anstis, S., and Kim, J. (2011). Local versus global perception of ambiguous motion displays. *J Vision* 11(3): 13, 1–12. Available at: http://www.journalofvision.org/content/11/3/13.

Avidan, G., Harel, M., Hendler, T., Ben-Bashat, D., Zohary, E., and Malach, R. (2002). Contrast sensitivity in human visual areas and its relationship to object recognition. *J Neurophysiol* 87: 3102–3116.

Baloch, A. A., and Grossberg, S. (1997). A neural model of high—level motion processing: line motion and form-motion dynamics. *Vision Res* 37(21): 3037–3059.

Baro, J. A., and Levinson, E. (1988). Apparent motion can be perceived between patterns with dissimilar spatial frequencies. *Vision Res* 28: 1311–1313.

Blair, C. B., Goold, J., Killebrew, K., & Caplovitz, G. P. (2013). Trackable features provide a cue to the angular velocity of rotating objects. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. doi: 10.1037/a0033055

Born, R. T., and Bradley, D. C. (2005). Structure and function of visual area MT. *Ann Rev Neurosci* 28: 157–189.

Braddick, O. J., O'Brien, J. M., Wattam-Bell, J., Atkinson, J., Hartley, T., and Turner, R. (2001). Brain areas sensitive to coherent visual motion. *Perception* 30: 61–72.

Braddick, O. J., O'Brien, J. M., Wattam-Bell, J., Atkinson, J., and Turner, R. (2000). Form and motion coherence activate independent but not dorsal/ventral segregated, networks in the human brain. *Curr Biol* 10: 731–734.

Bradley, D. C., Qian, N., and Andersen, R. A. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature* 373(6515): 609–611.

Brown, J. F. (1931). The visual perception of velocity. *Psychol Res* 14(1): 199–232.

Bruno, N., & Bertamini, M. (2013). Perceptual organization and the aperture problem. In J. Wagemans (Ed.), Oxford Handbook of Perceptual Organization (in press). Oxford, U.K.: Oxford University Press.

Burt, P., and Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychol Rev* 88: 171–195.

Campbell, F. W., & Maffei, L. (1979). Stopped visual motion. *Nature* 278: 192–193.

Campbell, F. W., & Maffei, L. (1981). The influence of spatial frequency and contrast on the perception of moving patterns. *Vision Res* 21: 713–721.

Caplovitz, G. P., Hsieh, P-J., & Tse, P. U. (2006). Mechanisms underlying the perceived angular velocity of a rigidly rotating object. *Vision Res* 46(18): 2877–2893.

Caplovitz, G. P., & Tse, P. U. (2007a). Rotating dotted ellipses: motion perception driven by grouped figural rather than local dot motion signals. *Vision Res* 47(15): 1979–1991.

Caplovitz, G. P., & Tse, P. U. (2007b). V3A processes contour curvature as a trackable feature for the perception of rotational motion. *Cerebral Cortex* 17(5): 1179–1189.

Caplovitz, G. P., Barroso, D. J., Hsieh, P. J., & Tse, P. U. (2008). fMRI reveals that non-local processing in ventral retinotopic cortex underlies perceptual grouping by temporal synchrony. *Hum Brain Map* 29(6): 651–661.

Caudek, C., Fantoni, C., & Domini, F. (2011). Bayesian modeling of perceived surface slant from actively-generated and passively-observed optic flow. *PLoS ONE* 6(4): 1–12.

Cavanagh, P., Arguin, M., and von Grünau, M. (1989). Interattribute apparent motion. *Vision Res* 29(9): 1197–1204.

Cavanagh, P., and Mather, G. (1989). Motion: the long and short of it. *Spatial Vis* 4: 103–129.

Dawson, M. R. W. (1991). The how and why of what went where in apparent motion: modeling solutions to the motion correspondence problem. *Psychol Rev* 33(4): 569–603.

Domini, F., & Caudek, C. (1999). Perceiving surface slant from deformation of optic flow. *J Exp Psychol Hum Percept Perform* 25(2): 426–444.

Domini, F., & Caudek, C. (2003). 3-D structure perceived from dynamic information: a new theory. *Trends Cogn Sci* 7(10): 444–449.

Domini, F., Caudek, C., & Tassinari, H. (2006). Stero and motion information are not independently processed by the visual system. *Vision Res* 46: 1707–1723.

Domini, F., Caudek, C., Turner, J., & Favretto, A. (1998). Discriminating constant from variable angular velocities in structure form motion. *Percept Psychophys* 60(5): 747–760.

Downing, P., and Treisman, A. (1995). The shooting line illusion: attention or apparent motion? *Invest Ophthalmol Vision Sci* 36: S856.

Downing, P., and Treisman, A. (1997). The line motion illusion: attention or impletion? *J Exp Psychol Hum Percept Perform* 23(3): 768–779.

Fantoni, C., Caudek, C., & Domini, F. (2010). Systematic distortions of perceived planar surface motion in active vision. *J Vision* 10(5): 12, 1–20.

Fantoni, C., Caudek, C., & Domini, F. (2012). Perceived slant is systematically biased in actively-generated optic flow. *PLoS ONE* 7(3): 1–12.

Faubert, J., and von Grünau, M. (1995). The influence of two spatially distinct primers and attribute priming on motion induction. *Vision Res* 35(22): 3119–3130.

Ferber, S., Humphrey, G. K. and Vilis, T. (2003). The lateral occipital complex subserves the perceptual persistence of motion-defined groupings. *Cereb Cortex* 13: 716–721.

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Curr Opin Neurobiol* 10: 438–443.

Foster, R., Fantoni, C., Caudeck, C., & Domini, F. (2011). Integration of disparity and velocity information for haptic and perceptual judgments of object depth. *Acta Psychol* 136: 300–310.

Georges, S., Seriès, P., Frégnac, Y., & Lorenceau, J. (2002). Orientation dependent modulation of apparent speed: Psychophysical evidence. *Vision Res* 42: 2757–2772.

Gepshtein, S., and Kubovy, M. (2000). The emergence of visual objects in spacetime. *Proc Natl Acad Sci USA* 97(14): 8186–8191.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gilaie-Dotan, S., Ullman, S., Kushnir, T., and Malach, R. (2001). Shape-selective stereo processing in human object- related visual areas. *Hum Brain Map* 15: 67–9.

Goodale, M., and Milner, A. (1992). Separate visual pathways for perception and action. *Trends Neurosci* 15: 20–25.

Green, B. F., Jr. (1961). Figure coherence in the kinetic depth effect. *J Exp Psychol* 62(3): 272–282.

Gilbert, C. D. (1992). Horizontal integration and cortical dynamics. *Neuron* 9: 1–13.

Gilbert, C. D. (1998). Adult cortical dynamics. *Physiol Rev* 78: 467–485.

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., and Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24: 187–203.

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., and Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21: 191–202.

Grill-Spector, K., Kourtzi, Z., and Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Res* 41: 1409–1422.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293(5539): 2425–2430.

Herzog, M. H., & Öğmen, H. (2013). Apparent motion and reference frames. In J. Wagemans (Ed.), Oxford Handbook of Perceptual Organization (in press). Oxford, U.K.: Oxford University Press.

Hikosaka, O., Miyauchi, S., and Shimojo, S. (1991). Focal visual attention produces motion sensation in lines. *Investig Ophthamol Vis Sci* 32(4): 176.

Hikosaka, O., Miyauchi, S., and Shimojo, S. (1993a). Focal visual attention produces illusory temporal order and motion sensation. *Vision Res* 33(9): 1219–1240.

Hikosaka, O., Miyauchi, S., and Shimojo, S. (1993b). Visual attention revealed by an illusion of motion. *Neurosci Res* 18(1): 11–18.

Hildreth, E. C., Ando, H., Andersen, R. A., and Treue, S. (1995). Recovering three-dimensional structure from motion with surface reconstruction. *Vision Res* 35(1): 117–137.

Hock, H. S. (2013). Dynamic grouping motion: A method for determining perceptual organization for objects with connected surfaces. In J. Wagemans (Ed.), Oxford Handbook of Perceptual Organization (in press). Oxford, U.K.: Oxford University Press.

Hsieh, P-J., Caplovitz, G. P., and Tse, P. U. (2005). Illusory rebound motion and the motion continuity heuristic. *Vision Res* 45(23): 2972–2985.

Hsieh, P-J., and Tse, P. U. (2006). Stimulus factors affecting illusory rebound motion. *Vision Res* 46(12): 1924–1933.

Hsieh, P-J., & Tse, P. U. (2007). Grouping inhibits motion fading by giving rise to virtual trackable features. *J Exp Psychol Hum Percept Perform* 33: 57–63.

Hubel, D. H., and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195: 215–243.

Kaiser, M. K. (1990). Angular velocity discrimination. *Percept Psychophys* 47: 149–156.

Kanizsa, G. (1951). Sulla polarizzazione del movimento gamma [The polarization of gamma movement]. *Arch Psichol Neurol Psichiatr* 3: 224–267.

Kanizsa, G. (1979). *Organization in Vision: Essays on Gestalt Perception*. New York: Praeger.

Kanwisher, N., Chun, M. M., McDermott, J., and Ledden, P. J. (1996). Functional imagining of human visual recognition. *Brain Res Cogn Brain Res* 5(1–2): 55–67.

Kenkel, F. (1913). Untersuchungen über den zusammenhang zwischen erscheinungsgrösse und erscheinungsbewegung bei einigen sogenannten optischen täuschungen. *Zeitschrift für Psychologie* 67: 358–449.

Kohler, P. J., Caplovitz, G. P., Hsieh, P-J., Sun, J., & Tse, P. U. (2010). Motion fading is driven by perceived, not actual angular velocity. *Vision Res* 50: 1086–1094.

Kohler, P. J., Caplovitz, G. P., & Tse, P. U. (2009). The whole moves less than the spin of its parts. *Attention, Percept Psychophys* 71(4): 675–679.

Kolers, P. A., and Pomerantz, J. R. (1971). Figural change in apparent motion. *J Exp Psychol* 87: 99–108.

Kolers, P. A., and von Grünau, M. (1976). Shape and color in apparent motion. *Vision Research* 16: 329-335.

Koyama, S., Sasaki, Y., Andersen, G. J., Tootell, R. B., Matsuura, M., and Watanabe, T. (2005). Separate processing of different global-motion structures in visual cortex is revealed by FMRI. *Curr Biol* 15(22): 2027–2032.

Kourtzi, Z., Erb, M., Grodd, W., and Bülthoff, H. H. (2003a). Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cereb Cortex* 13(9): 911–920.

Kourtzi, Z., and Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *J Neurosci* 20: 3310–3318.

Kourtzi, Z., and Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science* 293: 1506–1509.

Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., and Logothetis, N. K. (2003b). Integration of local features into global shapes. Monkey and human FMRI studies. *Neuron* 37(2): 333–346.

Lamme, V. A., Super, H., and Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Curr Opin Neurobiol* 8: 529–535.

Lichtenstein, M. (1963). Spatio-temporal factors in cessation of smooth apparent motion. *J Opt Soc Am* 53: 304–306.

Liu, T., and Cooper, L. A. (2003). Explicit and implicit memory for rotating objects. *J Exp Psychol Learn Mem Cogn* 29: 554–562.

Liu, T., Slotnick, S. D., and Yantis, S. (2004). Human MT+ mediates perceptual filling-in during apparent motion. *NeuroImage* 21(4): 1772–1780.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., and Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci* 92(18): 8135–8139.

McCarthy, J. D., Cordeiro, D., and Caplovitz, G. D. (2012). Local form-motion interactions influence global form perception. *Attention Percept Psychophys* 74: 816–823.

Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., and Tootell, R. B. H. (1999). The representation of real and illusory contours in human cortical visual areas revealed by fMRI. *J Neurosci* 19: 8560–8572.

Mirabella, G., and Norcia, A. N. (2008). Neural correlates of transformational apparent motion. *Perception* 37: 1368–1379.

Moore, C., and Engel, S. A. (2001). Neural response to perception of volume in the lateral occipital complex. *Neuron* 29: 277–286.

Moutoussis, K., Keliris, G., Kourtzi, Z., and Logothetis, N. (2005). A binocular rivalry study of motion perception in the human brain. *Vision Res* 45(17): 2231–2243.

Murray, S. O., Olshausen, B. A., and Woods, D. L. (2003). Processing shape, motion and three-dimensional shape-from-motion in the human cortex. *Cereb Cortex* 13: 508–516.

Nakayama, K., and Silverman, G. H. (1988b). The aperture problem II. Spatial integration of velocity information along contours. *Vision Res* 28(6): 747–753.

Navon, D. (1976). Irrelevance of figural identity for resolving ambiguities in apparent motion. *J Exp Psychol Hum Percept Perform* 2: 130–138.

Nowlan, S. J., and Sejnowski, T. J. (1995). A selection model for motion processing in area MT of primates. *J Neurosci* 15(2): 1195–1214.

Pack, C. C., and Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature* 409(6823): 1040–1042.

Pack, C. C., Gartland, A. J., and Born, R. T. (2004). Integration of contour and terminator signals in visual area MT of alert macaque. *J Neurosci* 24(13): 3268–3280.

Pack, C. C., Livingstone, M. S., Duffy, K. R., and Born, R. T. (2003). End-stopping and the aperture problem: two-dimensional motion signals in macaque V1. *Neuron* 39(4): 671–680.

Porter, K. B., Caplovitz, G. P., Kohler, P. J., Ackerman, C. M., & Tse, P. U. (2011). Rotational and translational motion interact independently with form. *Vision Res* 51: 2478–2487.

Ramachandran, V.S., Ginsburg, A. P., and Anstis, S. M. (1983). Low spatial frequencies dominate apparent motion. *Perception* 12: 457–461.

Ramachandran, V. S., and Gregory, R. L. (1978). Does colour provide an input to human motion perception? *Nature* 275: 55–56.

Schira, M. M., Fahle, M., Donner, T. H., Kraft, A., and Brandt, S. A. (2004). Differential contribution of early visual areas to the perceptual process of contour processing. *J Neurophysiol* 91(4): 1716–1721.

Seriès, P., Georges, S., Lorenceau, J., & Frégnac, Y. (2002). Orientation dependent modulation of apparent speed: a model based on the dynamics of feedforward and horizontal connectivity in V1 cortex. *Vision Res* 42: 2781–2797.

Spillmann, L., & De Weerd, P. (2003). Mechanisms of surface completion: perceptual filling-in of texture. In *Filling-in: From Perceptual Completion to Cortical Reorganization*, edited by L. Pessoa & P. De Weerd, pp. 81–105. Oxford: Oxford University Press.

Stelmach, L. B., and Herdman, C. M. (1991). Directed attention and perception of temporal order. *J Exp Psychol Hum Percept Perform* 17(2): 539–550.

Stelmach, L. B., Herdman, C. M., and McNeil, K. R. (1994). Attentional modulation of visual processes in motion perception. *Journal of Experimental Psychology: Human Perception and Performance* 20(1): 108-121.

Sternberg, S., and Knoll, R. L. (1973). The perception of temporal order: fundamental issues and a general model. In: *Attention and Performance*, Vol. IV, edited by S. Kornblum, pp. 629–685. New York: Academic Press.

Stone, J. V. (1999). Object recognition: view-specificity and motion-specificity. *Vision Res* 39: 4032–4044.

Stoner, G. R., and Albright, T. D. (1992). Motion coherency rules are form-cue invariant. *Vision Res* 32(3): 465–475.

Stoner, G. R., and Albright, T. D. (1996). The interpretation of visual motion: evidence for surface segmentation mechanisms. *Vision Res* 36(9): 1291–1310.

Titchener, E. B. (1908). *Lecture on the Elementary Psychology of Feeling and Attention*. New York: McMillan.

Tootell, R. B., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Liu, A. K., Reppas, J. B., Sereno, M. I., and Dale, A. M. (1997). Functional analysis of V3A and related areas in human visual cortex. *J Neurosci* 17(18): 7060–7078.

Tse, P. U. (2006). Neural correlates of transformational apparent motion. *NeuroImage* 31(2): 766–773.

Tse, P. U., and Caplovitz, G. P. (2006). Contour discontinuities subserve two types of form analysis that underlie motion processing. In: *Progress in Brain Research* 154: *Visual Perception. Part I. Fundamentals of Vision: Low and Mid-level Processes in Perception*, edited by S. Martinez-Conde, S. L. Macknick, L. M. Martinez, J-M. Alonso, and P. U. Tse, pp. 271–292. Amsterdam: Elsevier.

Tse, P. U., and Cavanagh, P. (1995). Line motion occurs after surface parsing. *Invest Ophth Vision Sci* 36: S417.

Tse, P. U., Cavanagh, P., and Nakayama, K. (1996). The roles of attention in shape change apparent motion. *Invest Ophthalmol Vision Sci* 37: S213.

Tse, P. U., Cavanagh, P., and Nakayama, K. (1998). The role of parsing in high-level motion processing. In: *High-Level Motion Processing: Computational, Neurobiological, and Psychophysical Perspectives*, edited by T. Watanabe, pp. 249–266. Cambridge, MA: MIT Press.

Tse, P. U., and Logothetis, N. K. (2002). The duration of 3-d form analysis in transformational apparent motion. *Percept Psychophys* 64(2): 244–265.

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

Ungerleider, L., and Mishkin, M. (1982). Two cortical visual systems. In: *Analysis of Visual Behavior*, edited by D. Ingle, M. Goodale, and R. Mansfield, pp. 549–586. Cambridge, MA: MIT Press.

Vaina, L. M., Gryzacz, N. M., Saiviroonporn, P., LeMay, M., Bienfang, D. C., and Conway, A. (2003). Can spatial and temporal motion integration compensate for deficits in local motion mechanisms? *Neuropsychologia* 41: 1817–1836.

Vanduffel, W., Fize, D., Peuskens, H., Denys, K., Sunaert, S., Todd, J. T., and Orban, G. A. (2002). Extracting 3D from motion: differences in human and monkey intraparietal cortex. *Science* 298: 413–415.

Vezzani, S., Kramer, P., & Bressan, P. (2013). Stereokinetic effect, kinetic depth effect, and structure from motion. In J. Wagemans (Ed.), Oxford Handbook of Perceptual Organization (in press). Oxford, U.K.: Oxford University Press.

Victor, J. D., and Conte, M. M. (1990). Motion mechanisms have only limited access to form information. *Vision Res* 30: 289–301.

von Grünau, M., and Faubert, J. (1994). Intraattribute and interattribute motion induction. *Perception* 23(8): 913–928.

von der Heydt, R., Peterhans, E., and Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science* 224(4654): 1260–1262.

Wallach, H. (1935). Uber visuell wahrgenommene Bewegungsrichtung. *Psychol Forsch* 20: 325–380.

Wallach, H., and O'Connell, D. N. (1953). The kinetic depth effect. *J Exp Psychol* 45(4): 205–217.

Wallach, H., Weisz, A., & Adams, P. A. (1956). Circles and derived figures in rotation. *Am J Psychol* 69: 48–59.

Wallis, G., and Bülthoff, H. (2001). Effects of temporal association on recognition memory. *Proc Natl Acad Sci USA* 98(8): 4800–4804.

Weiss, Y., & Adelson, E. H. (2000). Adventures with gelatinous ellipses—constraints on models of human motion analysis. *Perception* 29: 543–566.

Wertheimer, M. (1961). Experimental studies on the seeing of motion. In T. Shipley (Ed.), Classics in psychology (pp. 1032-1088). New York: Philosophical Library. (Original work published 1912)

Zhuo, Y., Zhou, T. G., Rao, H. Y., Wang, J. J., Meng, M., Chen, M., Zhou, C., and Chen, L. (2003). Contributions of the visual ventral pathway to long-range apparent motion. *Science* 299: 417–420.