Interactions between motion and form processing in the human visual system

George Mather[1]*, Andrea Pavan[2], Rosilari Bellacosa Marotti[3], Gianluca Campana[3] and Clara Casco[3]

[1] School of Psychology, University of Lincoln, Lincoln, UK
[2] Institut für Psychologie, Universität Regensburg, Regensburg, Germany
[3] Department of General Psychology, University of Padua, Padua, Italy

The predominant view of motion and form processing in the human visual system assumes that these two attributes are handled by separate and independent modules. Motion processing involves filtering by direction-selective sensors, followed by integration to solve the aperture problem. Form processing involves filtering by orientation-selective and size-selective receptive fields, followed by integration to encode object shape. It has long been known that motion signals can influence form processing in the well-known Gestalt principle of common fate; texture elements which share a common motion property are grouped into a single contour or texture region. However, recent research in psychophysics and neuroscience indicates that the influence of form signals on motion processing is more extensive than previously thought. First, the salience and apparent direction of moving lines depends on how the local orientation and direction of motion combine to match the receptive field properties of motion-selective neurons. Second, orientation signals generated by "motion-streaks" influence motion processing; motion sensitivity, apparent direction and adaptation are affected by simultaneously present orientation signals. Third, form signals generated by human body shape influence biological motion processing, as revealed by studies using point-light motion stimuli. Thus, form-motion integration seems to occur at several different levels of cortical processing, from V1 to STS.

Keywords: motion sensitivity, motion-streaks, motion perception, motion-form interactions, biological motion

INTRODUCTION

Anatomical and physiological studies of primates have identified over 50 distinct visual processing areas in the cerebral cortex (see Felleman and Van Essen, 1991). Ungerleider and Mishkin (1982) proposed that these multiple areas are organized into two major processing streams, known as the ventral stream and the dorsal stream, both originating in the primary visual cortex. This proposed division has since become widely established as a fundamental organizing principle in the primate visual system. The ventral stream travels into the temporal lobe, including cortical areas V4, TEO, and TE, and is thought to be crucial for the visual recognition of objects (also known as the "what" stream). The dorsal stream courses into the parietal cortex, and includes areas V3, MT, and MST, and is thought to be crucial for motion integration, for encoding spatial relationships between objects and for visual guidance toward objects (also known as the "where" stream). Single-unit recording studies are consistent with the two streams hypothesis. For example, neurons in areas forming part of the ventral stream show selectivity for color, shape and texture while those forming part of the dorsal stream show selectivity for the direction and speed of visual motion (see review in Maunsell and Newsome, 1987; Ungerleider and Pasternak, 2004).

The use of parallel streams to process different visual attributes has several merits (Marr, 1982). Modularity allows each stream to optimize its processing for the relevant visual attribute, rather than compromise for the sake of generality. For example, form processing is best served by high spatial acuity and low temporal acuity in order to code fine details reliably, while motion processing can sacrifice fine spatial acuity in favor of sensitivity to rapid temporal change. Moreover, modularity ensures that limitations or errors in processing output remain confined, rather than propagate across attributes.

However, in recent years evidence has accumulated which is inconsistent with the established principle of parallel, modular processing streams. The evidence reviewed in this paper demonstrates that form and motion are not processed independently in the visual system. On the contrary there is extensive interplay between form and motion processing systems which relies on a continuous exchange of information between different processing stages. The Gestalt psychologists, for example, recognized signs of this interaction long before the two streams hypothesis was proposed, when they formulated the principle of "common fate." An invisible form composed of randomly arranged dots against a dotted background becomes immediately visible as soon as it moves, by virtue of the common fate of its dots, which all move together with a common speed and direction (see Uttal et al., 2000; see also Ledgeway and Hess, 2002 for similar results with motion-defined spatial contours; and Edwards, 2009 for motion-form interactions in common fate). This kind of figure-ground segregation shows clearly that forms can emerge from motion

processing, in the absence of any other cue. The following sections review evidence for three other kinds of motion-form interaction, two at lower-levels of analysis and the other at higher-levels.

MOVING LINES

One form of interaction between orientation signals and motion signals occurs in early visual areas. There is extensive physiological evidence that the receptive fields of direction selective neurons at low levels of analysis (V1) extract the motion component orthogonal to local orientation (Hubel and Wiesel, 1968; De Valois et al., 1982), so their directional response is ambiguous (the "aperture problem"). Neurons in extrastriate cortex (MT) solve the problem by integrating the responses of different V1 cells (Simoncelli and Heeger, 1998). Motion-selective cells in V1 respond more strongly to retinal motion in the direction perpendicular to their preferred orientation than to other directions. This response to the orthogonal motion component may explain a variety of perceptual phenomena. For example, the perceived speed of a line is more veridical when oriented orthogonally to its direction than when the line is tilted (Castet et al., 1993; Scott-Brown and Heeley, 2001). Furthermore, when bars slanted slightly away from vertical are oscillated up and down, the trajectories of the bars quickly become influenced by their orientation (Tse and Hsieh, 2007); the bars are perceived as moving up and down, but also at the same time sideways, creating the impression that the bars are following an elliptical trajectory.

The salience of a moving target line, i.e., the observers' ability to segment it from background noise lines, also depends on its orientation. Salience is increased when the orientation of the target line and its direction of motion are appropriately combined to match the property of the receptive field tuned to the orthogonal motion component. Indeed, when this component is available, the orientation of the line (Casco et al., 2006) and its motion direction are more easily discriminated. This has been shown for both two-frame (Casco et al., 2001) and multi-frame-motion sequences (Pavan et al., 2011). This last result in particular agrees with Nakayama et al.'s (1985) suggestion that spatial integration of motion signals is most efficient in a direction orthogonal to orientation. In multi-frame displays Pavan et al. (2011) showed that the consistent velocity of the orthogonal motion component in a target line allowed observers to detect it in the presence of the random frame-to-frame velocity and direction of noise lines. On the other hand, collinearity between target and noise does not aid detection (Alberti et al., 2010).

Thus, although there are end-stopped neurons in V1 that respond to the motion of line-terminators independently of line orientation (Pack et al., 2003), the orthogonal motion component is nevertheless important, and has been shown to affect the response of motion-selective neurons at later stages in MT (Pack and Born, 2001). The orthogonal component generates motion signals that may hinder the perception of veridical motion (Tse and Hsieh, 2007), but it can also improve motion segregation and grouping at very early stages of visual processing (Alberti et al., 2010; Pavan et al., 2011).

The end-stopped neurons reported by Pack et al. (2003) may also be implicated in another kind of motion-form interaction involving moving lines, in which the apparent direction of the lines is influenced by the shape of the aperture through which they are viewed. When an obliquely oriented drifting grating is presented behind an elongated horizontal aperture, the grating bars appear to move horizontally along the long axis of the aperture rather than obliquely, perpendicular to their own orientation (the "barberpole effect"). The bar or line terminators at the edge of the aperture appear to be particularly important for determining apparent direction (see Lorenceau and Shiffrar, 1992; Kooi, 1993; Fisher and Zanker, 2001; Badcock et al., 2003; Edwards et al., 2013). Terminators are created by the spatial form of the stimulus window. Psychophysical and neurophysiological evidence from aperture effects caused by terminators indicates that the underlying motion-form interaction takes place in a cortical area normally associated with the dorsal stream, namely MT (Pack et al., 2003, 2004).

In sum, research on moving lines reveals complex interactions between orientation and motion direction at the earliest levels of cortical analysis up to the point at which the aperture problem is solved, demonstrating that processing of these two attributes is inextricably linked.

MOTION-STREAKS

Sensory neurons cannot respond instantaneously to sudden, impulsive stimuli such as flashes of light. Instead their response builds up to a peak and then dissipates over a period ranging from tens to hundreds of milliseconds. For example, the response of retinal cone photoreceptors shows a peak ~70 ms after a bright flash and a trough at 150 ms (biphasic response); rod (monophasic) response peaks at 200 ms after the flash and does not return to baseline until a further 300 ms have elapsed (Schnapf and Baylor, 1987). Consequently, when a stimulus element translates rapidly across the retina, it leaves behind a trail of waning neural activity that is the likely neural substrate of "persistence of vision"; the motion-streaks seen behind bright moving objects such as fireworks. Persistence of vision can be viewed as an undesirable consequence of neural responses because of the motion blur it creates, and the biphasic temporal response of cones in bright light is arguably an attempt by the visual system to minimize its impact. But one obvious property of motion-streaks is potentially useful during motion processing: they are bound to be aligned with the axis of motion.

Cells tuned the orientation of motion-streaks should be maximally activated by them. Thus, rapid retinal motion produces responses both in motion-selective cells tuned to that direction, and in orientation-selective cells tuned to an orientation aligned with the axis of the motion—the motion-streak. Psychophysical evidence from orientation detection and after-effects, as well as recent neuroimaging data, is consistent with the view that motion-streaks excite orientation-tuned cells in the human visual system (Alais et al., 2011; Apthorp et al., 2011, 2013). Geisler (1999) proposed that the outputs of motion- and orientation-selective cells are combined in visual cortex to create a "spatial motion-direction" (SMD) sensor tuned to both streak orientation and motion direction. He also presented psychophysical evidence for the existence of such sensors. Luminance detection thresholds were measured for moving Gaussian blobs, in the presence of dynamic random line masks oriented either parallel or

orthogonal to the axis of motion. Mask orientation had no effect on thresholds at low blob speeds, but above a critical speed, parallel masks elevated detection thresholds relative to orthogonal masks, consistent with the SMD sensor. A limitation of this experiment is that it did not specifically measure motion discrimination, but instead employed a 2AFC detection paradigm. So one cannot be sure that the masking effect revealed anything about motion perception.

Ross et al. (2000) generated static random Glass patterns by taking a field of randomly positioned dots, and giving each dot a partner displaced from it by a short distance corresponding to rotation of the original dot about the center of the pattern by a fixed angle. When a series of such uncorrelated patterns is presented rapidly, observers report apparent rotation even though there is no dot-to-dot correspondence between successive patterns. Ross et al. (2000) interpret this effect as consistent with Geisler's (1999) SMD sensor (see also Krekelberg et al., 2005 for similar results). Burr and Ross (2002) addressed the limitation of Geisler's original threshold study by employing a task that required observers to discriminate the direction of motion. Thresholds were higher for random line masks parallel to motion direction than for masks perpendicular to motion direction, consistent with Geisler's findings. Edwards and Crane (2007) further provided evidence of a motion-streak mechanism using a 3-frame global-motion stimulus and manipulating the strength of the motion-streak. When the same dots carried the global-motion signal over successive motion frames (long-streak condition) lower thresholds were obtained at high speeds (consistent with a motion-streak system). This facilitation decreased markedly at low contrast, due to reduced motion-streak magnitude and to the low contrast sensitivity of form cells contributing to motion-streak extraction. In addition to their effect on motion thresholds, motion-streaks also alter the appearance of supra-threshold motion. Several papers report changes in the apparent direction of moving elements when they are superimposed on a static background of tilted lines (see Swanston, 1984; Khuu, 2012). A possible mechanism for this direction effect involves mutual inhibition between orientation-selective cells, some of which are activated by the tilted background while others are activated by the motion-streak. The resulting angle-expansion effect propagates to the motion system via the SMD sensor.

On the basis of the research surveyed so far, it cannot be argued that form and motion are processed by completely independent systems. Evidence indicates that the interactions between orientation signals and motion signals are likely to occur in early visual areas (e.g., V1, V2).

Claims for motion-form interactions beyond V1/V2 cannot be based simply on evidence for long-range interactions, since these can occur in V1 as contextual modulation of responses (Alexander and van Leeuwen, 2010). Instead they should relate to effects associated with the specific functions performed by higher-level cortical areas. Area MT is believed to be involved in the integration of directional motion signals. For example, adaptation to two superimposed fields of dots moving in different directions normally produces a unidirectional motion after-effect (MAE) in the direction opposite to the vector average of the adapting directions (Mather, 1980; Verstraten et al., 1994; van der Smagt et al., 1999; Verstraten et al., 1999; von Grünau, 2002; Alais et al., 2005), and the integration of the two adapting motion components is thought to occur in extrastriate cortex in the dorsal stream, most likely in area MT as mentioned earlier. Mather et al. (2012) psychophysically investigated motion-form interactions at this integration stage of processing. Their results showed that superimposing a static grating orthogonal to the direction of the resultant unidirectional MAE during adaptation reduced the strength of the MAE relative to a condition in which the grating was parallel to the resultant MAE direction. Thus, the strength of bi-directional motion adaptation was modulated by simultaneously presented orientation signals. These findings provide evidence that form and motion signals interact at the global motion level where moving components are integrated, i.e., at a level of processing which is clearly part of the two-stream architecture.

Neurons in area MST of the dorsal stream are closely associated with the analysis of global patterns of motion (i.e., optic flow; Graziano et al., 1994). Neurons in the dorsal part of area MST (i.e., MSTd) of the macaque have large receptive fields (from 10° up to 100°; Desimone and Ungerleider, 1986; Tanaka and Saito, 1989) and show selectivity to optic flow and to its components (Sakata et al., 1985, 1986; Saito et al., 1986; Tanaka et al., 1986, 1989; Tanaka and Saito, 1989; Duffy and Wurtz, 1991b; Lagae et al., 1993; Graziano et al., 1994). There is psychophysical evidence for motion adaptation at the level of optic flow analysis, in the form of the phantom MAE. In phantom MAEs, adaptation of some parts/sectors of the visual field to complex motion components such as expansion (or contraction) induces the perception of contraction (or expansion) in other (non-adapted) parts of the visual field during testing. The phantom MAE is likely to reflect adaptation of cells with large complex receptive fields at the level of MST (Regan and Beverly, 1985; Desimone and Ungerleider, 1986; Tanaka and Saito, 1989; Duffy and Wurtz, 1991a; Lagae et al., 1993; Graziano et al., 1994; Morrone et al., 1995; Snowden and Milne, 1996, 1997; Burr et al., 1998). Pavan et al. (2013) used the phantom MAE to test for the presence of form-motion interactions at this high-level site of adaptation in the dorsal stream. Their results showed that adding a concentric grating orthogonal to radial optic flow during adaptation suppressed the duration of the phantom MAE, compared to a radial grating parallel to the global pattern of motion. This may indicate an interaction between form and motion signals at the level in which optic flow is processed.

Recent evidence indicates that inferences about stimulus selectivity based on an adaptation paradigm are not necessarily straightforward (Rentzeperis et al., 2012). Nevertheless, in the case of motion-form interactions during optic flow analysis, evidence from Niehorster et al.'s (2010) discrimination study bears out Pavan et al.'s (2013) adaptation study. Niehorster et al. (2010) showed that human heading perception in a heading direction discrimination task was based on a combination of motion (optic flow component) and form (radial glass patterns) signals. There is evidence for neurons in the form processing stream which are sensitive to these radial streak patterns (Gallant et al., 1993; Ostwald et al., 2008). The visual system may take advantage of

the close correspondence between visual form and motion signals generated by locomotion, combining the two during high-level optic flow processing.

BIOLOGICAL MOTION

Johansson (1973) introduced highly impoverished "point-light walker" movies in which moving human figures are visible only by means of isolated points of light fixed at the major joints (ankles, knees, hips, wrists, elbows, shoulders). Naive observers are able rapidly and reliably to perceive many human attributes in these movies despite the paucity of available information, including the actor's gender, mood, and action type (see review in Blake and Shiffrar, 2007). Point-light walker displays are now also widely known as biological motion displays. In the forty years since their introduction biological motion displays have attracted debate and dispute regarding the neural processes which mediate their perception; do they involve form analysis (the ventral stream), or motion analysis (the dorsal stream), or both? At first sight one might think that biological motion displays specifically target motion analysing processes, since there are no explicit visual connections between any of the dots. Indeed many psychophysical studies attest to the importance of motion signals. Spatiotemporal properties such as display duration, dot displacement distance and inter-frame interval are all critical to biological motion perception, consistent with a reliance on information in the dorsal stream (e.g., Johansson, 1976; Mather et al., 1992; Thornton et al., 1998). However, low-pass spatial filtering of any single frame in a biological motion sequence would reveal a blurred, body-shaped form which could serve as a stimulus for form processing. A number of spatial properties do affect biological motion perception in a way that implicates processes in the ventral stream. Beintema and colleagues limited the display lifetime of individual dots (Beintema et al., 2006) or shifted dots around the body on a frame-by-frame basis so that they were not placed consistently at the joints (Beintema and Lappe, 2002), and at least some degree of biological motion perception survived both manipulations. Thus, it is difficult to argue against the proposition that biological motion analysis involves both the dorsal *and* ventral streams. The question then arises as to where in the cortex is the information from the two streams combined. Regions within the rostral Superior Temporal Sulcus (STS) receive information from both streams, so STS is a likely area of convergence (Ungerleider and Pasternak, 2004). Geise and colleagues have developed and tested a computational model of biological motion analysis that conforms to this architecture: separate analyses in the dorsal and ventral streams converge on a common representation in high-level areas such as STS (Giese and Poggio, 2003; Fleischer and Giese, 2012). Neuroimaging data is consistent with this hierarchy, and also implicates extrastriate and fusiform body areas (EBA and FBA; Jastorff and Orban, 2009). Fleischer and Giese (2012) acknowledge, however, that segregation of signals until they reach very late stages of cortical analysis may be an oversimplification. Many studies use background "noise" dots to mask form-based cues, either moving in a random fashion or in a way that mimics the local motion of the figure dots. The presumption is that noise dots abolish form cues, since the form is invisible in each frame.

However, given the well-known common fate Gestalt principle described in the Introduction, one could argue in favor of a low-level inter-play between form and motion processing in which motion-mediated common fate allows the visual system to segregate dots representing the body form from the background dots, and later motion and form processes extract gender, mood and so on.

Form processing of biological motion in Giese and Poggio's (2003) model includes "snap-shot" neurons which are selective for specific body shapes that are adopted during movement. The output of these ventral stream neurons allows motion to be inferred from body shape. As Giese and Poggio (2003) state (p. 184) "active snapshot neurons pre-excite neurons that encode temporally subsequent configurations, and inhibit neurons that encode other configurations." Lange and Lappe's (2006) form-based model of biological motion analysis employs similar posture-specific form cells to encode the different body configurations adopted while walking, and a coding scheme based on their sequential activation.

Artists have traditionally been able to convey an impression of motion in static artworks such as painting and sculpture using poses which imply motion because they would be physically impossible for a human actor to hold for any length of time. Vision scientists call such static depictions of action "implied motion." The snap-shot or posture-specific neurons in the ventral stream proposed by Giese and Poggio (2003) and Lange and Lappe (2006) are a plausible neural substrate for the encoding of implied motion. There is accumulating evidence that activity originating in such neurons finds it way to cells in the dorsal stream. Senior et al. (2000) used fMRI to identify brain regions activated by video clips of objects in motion, and clips of the same objects at rest. Activation in dorsal area MT was higher while participants viewed the movie clips, as one would expect for an area involved in motion analysis. Interestingly, Senior et al. (2000) also found higher activation in MT while participants viewed still images implying motion, compared to images containing no implied motion. Similar results were reported by Kourtzi and Kanwisher (2000). A plausible source of MT activation by implied motion is cells in the dorsal stream that are sensitive to the motion implied by form; snap-shot neurons. Alternatively, recent neuroimaging results indicate that cells sensitive to static body shape and to motion are actually intermingled in area MT (Ferri et al., 2013). Thus, the interaction between the form and motion pathways may not be confined to convergence at the level of STS, but could involve cross-activation at the level of MT. Winawer et al. (2008) exposed experimental participants to rapidly presented sequences of unrelated static images each containing implied motion, and reported that this "adaptation" generated a motion aftereffect on a directionally ambiguous dynamic test pattern (see also Pavan et al., 2011; Pavan and Baggio, 2013, for similar results). Such results would be consistent with cross-activation of MT by neurons in the dorsal stream, because MT neurons have long been associated with motion adaptation. However, Morgan et al. (2012) sound a note of caution, arguing that the post-adaptation directional bias found by Winawer et al. (2008) could be due to a shift in decision bias rather than a shift in the relative activity of direction-selective neurons.

Unlike Winawer et al. (2008), Pavan et al. (2011) employed a control adapting condition that did not contain implied motion, but still allowed the possibility of response bias. They did not obtain an after-effect in this condition.

SUMMARY

Visual motion and form information is inextricably linked in the sense that motion is, by definition, spatiotemporal; change over both time *and* space. The research reviewed here indicates that the two components of motion interact at multiple levels of processing. Prior to segregation into parallel dorsal and ventral streams, the salience and apparent direction of moving lines depends jointly on line orientation and motion. The SMD sensor is designed to exploit the orientation signals generated by fast motion in the form of motion-streaks. Evidence from research on implied motion and biological motion indicates that interactions between form and motion processes also occur after the point at which the dorsal and ventral streams diverge, probably in area MT, as well as at the point of convergence in STS.

Thus, the visual system seems to take advantage both of modular processing and data sharing, by allowing data to flow between specialized neural processing streams. The theoretical justification for these interactions rests on the high degree of correlation between the signals in different modules, due to their common origin in natural images. Integration of signals across processing modules serves to minimize signal redundancy and maximize signal reliability.

ACKNOWLEDGMENTS

REFERENCES

Alais, D., Apthorp, D., Karmann, A., and Cass, J. (2011). Temporal integration of movement: the time-course of motion streaks revealed by masking. *PloS ONE* 6:e28675. doi: 10.1371/journal.pone.0028675

Alais, D., Verstraten, F. A., and Burr, D. C. (2005). The motion aftereffect of transparent motion: two temporal channels account for perceived direction. *Vision Res.* 45, 403–412.

Alberti, C., Pavan, A., Casco, C., and Campana, G. (2010). Segmentation by single and combined features involves different contextual influences. *Vision Res.* 50, 1065–1073.

Alexander, D. M., and van Leeuwen, C. (2010). Mapping of contextual modulation in the population response of primary visual cortex. *Cogn. Neurodyn.* 4, 1–24.

Apthorp, D., Cass, J., and Alais, D. (2011). The spatial tuning of "motion streak" mechanisms revealed by masking and adaptation. *J. Vis.* 11, 1–16.

Apthorp, D., Schwarzkopf, D. S., Kaul, C., Bahrami, B., Alais, D., and Rees, G. (2013). Direct evidence for encoding of motion streaks in human visual cortex. *Proc. R. Soc. B Biol. Sci.* 280, 1471–2954.

Badcock, D. R., McKendrick, A. M., and Ma-Wyatt, A. (2003). Pattern cues disambiguate perceived direction in simple moving stimuli. *Vision Res.* 43, 2290–2301.

Beintema, J. A., Georg, K., and Lappe, M. (2006). Perception of biological motion from limited lifetime stimuli. *Percept. Psychophys.* 68, 613–624.

Beintema, J. A., and Lappe, M. (2002). Perception of biological motion without local image motion. *Proc. Natl. Acad. Sci.* 99, 5661–5663.

Blake, R., and Shiffrar, M. (2007). Perception of human motion. *Annu. Rev. Psychol.* 58, 47–73.

Burr, D. C., Morrone, M. C., and Vaina, L. M. (1998). Large receptive fields for optic flow detection in humans. *Vision Res.* 38, 1731–1743.

Burr, D. C., and Ross, J. (2002). Direct evidence that speedlines influence motion mechanisms. *J. Neurosci.* 22, 8661–8664.

Casco, C., Caputo, G., and Grieco, A. (2001). Discrimination of an orientation difference in dynamic textures. *Vision Res.* 41, 275–284.

Casco, C., Grieco, A., Giora, E., and Martinelli, M. (2006). Saliency from orthogonal velocity component in texture segregation. *Vision Res.* 46, 1091–1098.

Castet, E., Lorenceau, J., Shiffrar, M., and Bonnet, C. (1993). Perceived speed of moving lines depends on orientation, length, speed and luminance. *Vision Res.* 33, 1921–1936.

Desimone, R., and Ungerleider, L. G. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J. Comp. Neurol.* 248, 164–189.

De Valois, R. L., Yund, E. W., and Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res.* 22, 531–544.

Duffy, C. J., and Wurtz, R. H. (1991a). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J. Neurophysiol.* 65, 1329–1345.

Duffy, C. J., and Wurtz, R. H. (1991b). Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. *J. Neurophysiol.* 65, 1346–1359.

Edwards, M., Cassanello, C. R., Badcock, D. R., and Nishida, S. (2013). Effect of form cues on 1D and 2D motion pooling. *Vision Res.* 76, 94–104.

Edwards, M. (2009). Common-fate motion processing: Interaction of the On and Off pathways. *Vision Res.* 49, 429–438.

Edwards, M., and Crane, M. F. (2007). Motion streaks improve motion detection. *Vision Res.* 47, 828–833.

Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cortex. *Cereb. Cortex* 1, 1–47.

Ferri, S., Kolster, H., Jastorff, J., and Orban, G. A. (2013). The overlap of the EBA and the MT/V5 cluster. *Neuroimage* 66, 412–425.

Fisher, N., and Zanker, J. M. (2001). The directional tuning of the barber-pole illusion. *Perception* 30, 1321–1336.

Fleischer, F., and Giese, M. A. (2012). "Computational mechanisms of the visual processing of action stimuli," in *People Watching: Social, Perceptual and Neurophysiological Studies of Body Perception,* eds K. L. Johnson and M. Shiffrar (Oxford: Oxford University Press), 388–413.

Gallant, J. L., Braun, J., and Van Essen, D. C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* 259, 100–103.

Geisler, W. S. (1999). Motion streaks provide a spatial code formation direction. *Nature* 400, 65–69.

Giese, M. A., and Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nat. Rev. Neurosci.* 4, 179–192.

Graziano, M. S. A., Andersen, R. A., and Snowden, R. J. (1994). Tuning of MST neurons to spiral stimuli. *J. Neurosci.* 14, 54–67.

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

Jastorff, J., and Orban, G. A. (2009). Human functional magnetic resonance imaging reveals separation and integration of shape and motion cues in biological motion processing. *J. Neurosci.* 29, 7315–7329.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Percep. Psychophys.* 14, 201–211.

Johansson, G. (1976). Spatio-temporal differentiation and integration in visual-motor perception. *Psychol. Res.* 38, 379–393.

Khuu, S. K. (2012). The role of motion streaks in the perception of the kinetic Zollner illusion. *J. Vis.* 12, 1–14.

Kooi, F. L. (1993). Local direction of edge motion causes and abolishes the barberpole illusion. *Vision Res.* 33, 2347–2351.

Kourtzi, Z., and Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *J. Cogn. Neurosci.* 12, 48–55.

Krekelberg, B., Vatakis, A., and Kourtzi, Z. (2005). Implied motion from form in the human visual cortex. *J. Neurophysiol.* 94, 4373–4386.

Lange, J., and Lappe, M. (2006). A model of biological motion perception from configural form cues. *J. Neurosci.* 26, 2894–2906.

Lagae, L., Raiguel, S., and Orban, G. A. (1993). Speed and direction selectivity of macaque middle temporal neurons. *J. Neurophysiol.* 69, 19–39.

Ledgeway, T., and Hess, R. F. (2002). Rules for combining the outputs of local motion detectors to define simple contours. *Vision Res.* 42, 653–659.

Lorenceau, J., and Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Res.* 32, 263–273.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* New York, NY: Freeman.

Mather, G. (1980). The movement after-effect and a distribution shift model of direction coding. *Perception* 9, 379–392.

Mather, G., Pavan, A., Bellacosa, R. M., and Casco, C. (2012). Psychophysical evidence for interactions between visual motion and form processing at the level of motion integrating receptive fields. *Neuropsychologia* 50, 153–159.

Mather, G., Radford, K., and West, S. (1992). Low-level visual processing of biological motion. *Proc. R. Soc. London B* 249, 149–155.

Maunsell, J. H. R., and Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annu. Rev. Neurosci.* 10, 363–402.

Morgan, M. J., Dillenburger, B., Raphael, S., and Solomon, J. A. (2012). Observers can voluntarily shift their psychometric functions without losing sensitivity. *Attent. Percept. Psychophys.* 74, 185–193.

Morrone, M. C., Burr, D. C., and Viana, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature* 376, 507–509.

Nakayama, K., Silverman, G. H., MacLeod, D. I., and Mulligan, J. (1985). Sensitivity to shearing and compressive motion in random dots. *Perception* 14, 225–238.

Niehorster, D. C., Cheng, J. C., and Li, L. (2010). Optimal combination of form and motion cues in human heading perception. *J. Vis.* 10, 1–15.

Ostwald, D., Lam, J. M., Li, S., and Kourtzi, Z. (2008). Neural coding of global form in the human visual cortex. *J. Neurophysiol.* 99, 2456–2469.

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, 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, 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, 671–680.

Pavan, A., and Baggio, G. (2013). Linguistic representations of motion do not depend on the visual motion system. *Psychol. Sci.* 24, 181–188.

Pavan, A., Bellacosa, R. M., and Mather, G. (2013). Motion-form interactions beyond the motion integration level: evidence for interactions between orientation and optic flow signals. *J. Vis.* 13. (in press).

Pavan, A., Casco, C., Mather, G., Bellacosa, R. M., Cuturi, L. F., and Campana, G. (2011). The effect of spatial orientation on detecting motion trajectories in noise. *Vision Res.* 51, 2077–2084.

Pavan, A., Cuturi, L. F., Maniglia, M., Casco, C., and Campana, G. (2011). Implied motion from static photographs influences the perceived position of stationary objects. *Vision Res.* 51, 187–194.

Regan, D., and Beverly, K. I. (1985). Visual responses to vorticity and the neural analysis of optic flow. *J. Opt. Soc. Am. A* 2, 280–283.

Rentzeperis, I., Nikolaev, A. R., Kiper, D. C., and van Leeuwen, C. (2012). Relationship between neural response and adaptation selectivity to form and color: an ERP study. *Front. Hum. Neurosci.* 6:89. doi: 10.3389/fnhum.2012.00089

Ross, J., Badcock, D. R., and Hayes, A. (2000). Coherent global motion in the absence of coherent velocity signals. *Curr. Biol.* 10, 679–682.

Saito, H.-A., Yukie, M., Tanaka, K., Hikosaka, K., and Fukada, Y., and Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J. Neurosci.* 6, 145–157.

Sakata, H., Shibutani, H., Ito, Y., and Tsurugai, K. (1986). Parietal cortical neurons responding to rotary movement of visual stimulus in space. *Exp. Brain Res.* 61, 658–663.

Sakata, H., Shibutani, H., Kawano, K., and Harrington, T. L. (1985). Neural mechanisms of space vision in the parietal association cortex of the monkey. *Vision Res.* 25, 453–463.

Schnapf, J. L., and Baylor, D. A. (1987). How photoreceptor cells respond to light. *Sci. Am.* 256, 40–47.

Scott-Brown, K. C., and Heeley, D. W. (2001). The effect of the spatial arrangement of target lines on perceived speed. *Vision Res.* 41, 1669–1682.

Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., et al. (2000). The functional neuroanatomy of implicit-motion perception or "representational momentum". *Curr. Biol.* 10, 16–22.

Simoncelli, E., and Heeger, D. (1998). A model of neuronal responses in visual area MT. *Vision Res.* 38, 743–761.

Snowden, R. J., and Milne, A. B. (1996). The effects of adapting to complex motions: position invariance and tuning to spiral motions. *J. Cogn. Neurosci.* 8, 435–452.

Snowden, R. J., and Milne, A. B. (1997). Phantom motion after effects–evidence of detectors for the analysis of optic flow. *Curr. Biol.* 7, 717–722.

Swanston, M. T. (1984). Displacement of the path of perceived movement by intersection with static contours. *Percept. Psychophys.* 36, 324–328.

Tanaka, K., Fukada, Y., and Saito, H. (1989). Underlying mechanisms of the response specificity of the expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J. Neurophysiol.* 62, 642–656.

Tanaka, K., Hirosaka, K., Saito, H., Yukie, M., Fukuda, Y., and Iwai, E. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J. Neurosci.* 6, 134–144.

Tanaka, K., and Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *J. Neurophysiol.* 62, 626–641.

Thornton, I. M., Pinto, J., and Shiffrar, M. (1998). The visual perception of human locomotion. *Cogn. Neuropsychol.* 15, 535–552.

Tse, P. U., and Hsieh, P. J. (2007). Component and intrinsic motion integrate in 'dancing bar' illusion. *Biol. Cybern.* 96, 1–8.

Ungerleider, L. G., and Mishkin, M. (1982). "Two cortical visual systems," in *Analysis of Visual Behavior*, eds D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield (Cambridge, MA: MIT Press), 549–586.

Ungerleider, L. G., and Pasternak, T. (2004). "Ventral and dorsal cortical processing streams," in *The Visual Neurosciences* eds L. M. Chalupa and J. S. Werner (Cambridge, MA: MIT Press), 541–562.

Uttal, W. R., Spillmann, L., Stürzel, F., and Sekuler, A. B. (2000). Motion and shape in common fate. *Vision Res.* 40, 301–310.

van der Smagt, M. J., Verstraten, F. A., and van de Grind, W. A. (1999). A new transparent motion aftereffect. *Nat. Neurosci.* 2, 595–596.

Verstraten, F. A., Fredericksen, R. E., and van de Grind, W. A. (1994). Movement aftereffect of bi-vectorial transparent motion. *Vision Res.* 34, 349–358.

Verstraten, F. A., van der Smagt, M. J., Fredericksen, R., and van de Grind, E. W. A. (1999). Integration after adaptation to transparent motion: static and dynamic test patterns result in different aftereffect directions. *Vision Res.* 39, 803–810.

von Grünau, M. W. (2002). Bivectorial transparent stimuli simultaneously adapt mechanisms at different levels of the motion pathway. *Vision Res.* 42, 577–587.

Winawer, J., Huk, A. C., and Boroditsky, L. (2008). A motion aftereffect from still photographs depicting motion. *Psychol. Sci.* 19, 276–283.