1 Introduction

Accurate motion perception is essential for visually guided movement; complex behaviors such as chasing prey or catching a thrown ball require that an organism be able to rapidly determine the position and velocity of a moving object, and to anticipate its trajectory through space. Psychological research on motion perception has established that local and global motion are processed by separate mechanisms, an idea that dates back to Wertheimer's phenomenological distinction between fast "phi" and slower "beta" motion (Steinman et al., 2000). It is believed that local (or short range) mechanisms respond to the motion of visual features defined by luminance, or by motion energy in the Fourier domain (Adelson and Bergen, 1985), at short temporal and spatial scales.

It is thought that the first-order motion mechanism originates with the direction selective response of cells in V1. Classic models of local motion include local, linear filters that extract directional signals from space-time correlations in luminance contrast in a small region of space, in a way similar to what has been observed to drive the activity of cells in V1 . Psychophysical data are consistent with the notion that the size of the local motion filters are similar in size of the classical receptive field of a V1 neuron.

However, the perception of motion does not necessarily require features to differ in mean luminance from the background, does not necessarily require motion energy in the Fourier domain. Moving stimuli can be constructed that would not be consistently detectable to first-order mechanisms, yet nonetheless elicit strong perceptions of movement. These higher-order stimuli can be constructed variously by modulations in contrast, texture, or other stimulus features, but generally involve the change in position, over time, of some feature in the image (Lu and Sperling, 1995). There is evidence that global motion may have a different neural substrate from local motion. For example, adding higher-order motion to a display does not appear to change the threshold of detection for first-order motion Edwards and Badcock (1995); Cassanello et al. (2011), and there is some neuropsychological evidence of while a combination of linear spatiotemporal filters and output nonlinearities is successful at predicting the detection of first order motion, and corresponds to neural response properties in V1, stimuli can also be designed to be invisible to such a mechanism. Higher order motion stimuli can be created these stimuli nonetheless elicit a strong sensation of motion can also be created without any Fourier motion signal, by using stimuli that modulate contrast, texture, or other features - such stimuli being invisible to a system that relies on linear filter ; Another difference between long-range and short-range motion is that the former seems to be capable of tracking a feature object over distances larger than what can be achieved through individual local filters. Interestingly, the anatomical substrate of higher-order motion mechanism is still unclear. While area MT and MST have been proposed as candidates, response properties of cells in those regions do not appear to explain our sensations of higher order motion (Ilg and Churan, 2004).

One possible reason for having multiple motion estimation systems is that

local motion signals are not always a reliable indication of the global motion. In a complicated visual world, motion can come from many sources, and accurate perception of the movement of objects requires disambiguating those motion signals attributable to the object from irrelevant motions in the background or of other objects. Consider the task of trying to track the movement of a zebra among a background of waving grass. While many of the local signals will correctly signal the motion of the zebra, the motion of the grass creates a subset of local motion signals that are substantially incorrect. significant proportion of these individual receptive fields provide an incorrect signal , and a simple pooling of their signals over space might also be incorrect. To track the zebra requires integrating local motion signals over space and time in a way that is consistent with the global change in position of an object, while discarding local signals that are inconsistent with the object's trajectory.

Figure 1 shows an example of stimuli that combine a first-order motion component with an independent higher-order (position-tracking) component.

Figure 1 demonstrates how first order and higher order motion can interact. The movie shows two wheels, each composed of five moving elements. The elements are Gabor-like stimuli that window a moving carrier inside of a spatial envelope. The envelope moves independently of the carrier, so that the carrier produces first order motion while the envelope produces higher order motion. On the right side of the display, the local and global motion components move in the same direction; on the left side the local and global components are in opposite directions. Full details of the construction of this display are given in section **??**.

When the elements are viewed in isolation, as in Figure 1, the appearance of the motion of the elements is consistent with the motion of the envelope. The carrier motion causes a change in the perception of "smoothness" but does not strongly affect the apparent direction or even the apparent speed of the motion. However, when multiple elements are placed in proximity, but not overlapping, as in Figure 1, the apparent motion depends on whether the stimulus is viewed centrally or peripherally. When the five-element stimulus on the left is viewed centrally, the apparent direction of motion is consistent with the envelope. When the ring is viewed in the periphery, the apparent direction of motion matches that of the carrier. If the observer keeps attention on that stimulus while moving the eyes so as to move the stimulus from central to peripheral vision it is possible to make the apparent motion reverse.

From this demonstration it appears that having more than one target in a given region of space affects how first-order and higher-order motion are combined. That the appearance changes with retinal eccentricity of the stimulus suggests that the range of spatial interaction scales with the retinal eccentricity. A plausible explanation could be that the presence of flanking objects limits the ability to see movement of the envelope, allowing the carrier motion to determine the percept; that is, 'crowding' may be affecting howfirst-order and higher-order motion are combined.

In this paper we examine how first-order and higher order mechanisms interact in forming an overall perception of motion. In Experiment 1 we quantify
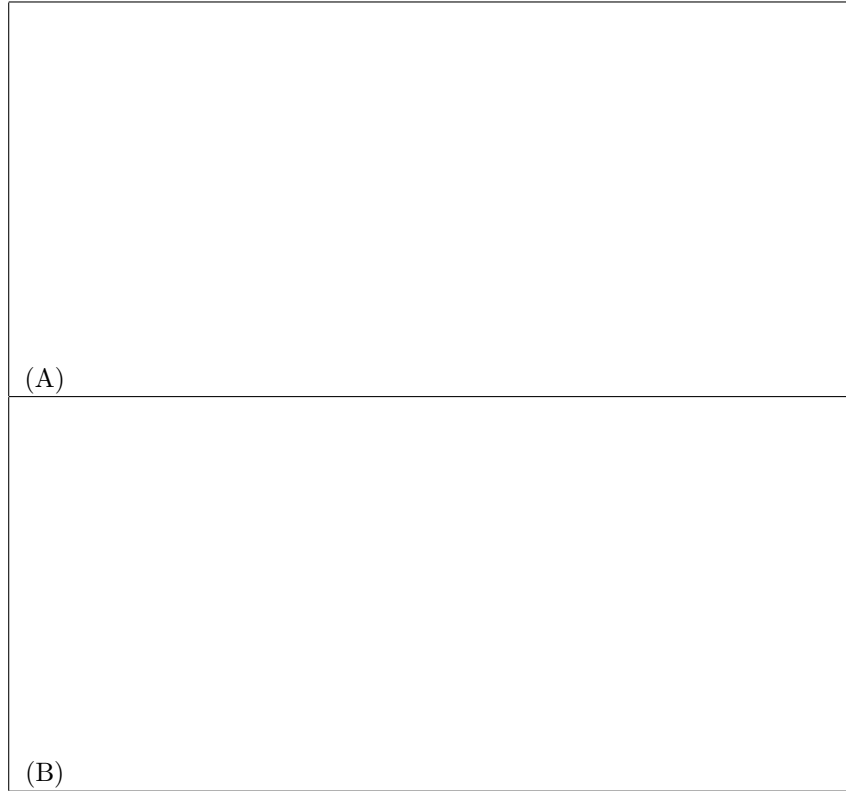
Figure 1: (A) Single motion elements, moving alone. At left, the carrier motion is opposite the envelope motion. At right, the carrier motion is in the same direction as the envelope motion. (B) Five elements, each identical to the single element in A, distributed around each fixation point. When elements are closely spaced, there is an eccentricity-dependent change in appearance. The appearance of the wheel on the left (with opposing carrier and envelope motions) changes depending on where the eyes fixate.

how element spacing or 'crowding' affects the apparent direction of motion and present a simple mixture model that assumes that global and local motion signals are combined In Experiment 2 we vary the number of elemnents independently of the spacing of targets and uncover a separate effect due to the number of targets on the screen.

References

Adelson, E. H. and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A*, 2(2):284–299.

Cassanello, C. R., Edwards, M., Badcock, D. R., and Nishida, S. (2011). No interaction of first- and second-order signals in the extraction of global-motion and optic-flow. *Vision Res*, 51(3):352–61.

Edwards, M. and Badcock, D. R. (1995). Global motion perception: no interaction between the first- and second-order motion pathways. *Vision Res*, 35(18):2589–602.

Ilg, U. J. and Churan, J. (2004). Motion perception without explicit activity in areas mt and mst. *Journal of Neurophysiology*, 92(3):1512–1523.

Lu, Z. L. and Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Res*, 35(19):2697–2722.

Steinman, R. M., Pizlo, Z., and Pizlo, F. J. (2000). Phi is not beta, and why Wertheimer's discovery launched the gestalt revolution. *Vision Research*, 40(17):2257–2264.