

# A FIXED CORTICAL DISTANCE GOVERNS PERCEPTUAL INTEGRATION OF LOCAL MOTION AND POSITION

PETER B. MEILSTRUP AND MICHAEL N. SHADLEN

## 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. Long-range or global motion does not necessarily require features to differ in mean luminance from the background, does not necessarily contain motion energy in the Fourier domain, and is characterized by operating over longer spatial and temporal scales than the short-range process. Strong global motion percepts can be elicited by stimuli that modulate in contrast or texture rather than in luminance. In general, global motion seems to involve the change in position, over time, of some feature in the image. [Lu and Sperling, 1995].

It is thought that local motion mechanism originates in cells in V1. Classic models of local motion include local 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. Global motion on the other hand consists of processes that track an object over distances larger than what can be achieved through individual local filters. A strong sensation of global motion can be created without any local motion signal, so it is possible that local and global mechanisms have independent origins. The anatomical substrate of global motion mechanism is unknown. While MT and MST have been proposed as candidates, *certain properties of these regions don't match what is known about our sensations of global motion. (Movshon, etc.)*

I have a distaste for 'global' as a name for the second process. It leads people down a wrong path of thinking about the wheel stimulus as having a 'global' movement i.e. globally consistent with solid body rotation of the wheel, whereas what we are *actually talking about here is a perception of the shift in position of different elements; it's quite easy to have a perception of the wheel stimulus that is not coherent with a solid body.*

cites:(Reichardt; Adelson; Simoncelli; Ullman & Marr; others)

cites

One thing that makes accurate estimation of motion particularly difficult is that the 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. Take the example of a frog trying to catch a zebra butterfly within waving grass. While many of the local signals will correctly signal the motion of the butterfly, the motion of the grass creates a subset of local motion signals that are substantially incorrect. Thus, a simple average of the motion over any receptive field large enough to cover more than a tiny portion of the scene will not accomplish the goal of tracking the motion of the butterfly. On the other hand, an average of motion signals within very small regions will also be incorrect; small regions containing only one edge cannot uniquely determine the motion of an object (the ‘aperture problem’, etc.) To track the butterfly, then 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.

In this paper we begin to examine how local motion and global changes in position interact in forming an overall perception of motion. We constructed stimuli that combined a local motion with global position shift; such that the local and the global motion can either be congruent or incongruent.<sup>1</sup> shows an example of how local and global motion signals can interact.

The movie shows two wheels, each composed of five moving elements. The elements are composed of a succession of brief wavelet pulses, at 100 ms intervals. These pulses have the same mean luminance as the background, and the envelope of each pulse does not move, however, the peaks and troughs of the wavelet move within the envelope during the pulse, giving the *local* component of the motion stimulus. The *global* motion component of the stimulus is created by offsetting the position of each pulse relative to the last. In this way we can control local and global motion of an object largely independently. 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 subsection 2.3.

This display elicits an illusion of motion reversal. The rotation of the spots around the fixation point on the right side of the display appears constant regardless of viewing angle. In contrast, the motion of the spots on the left side appears to change direction based on where one fixates. When fixating in the center of the left circle, the spots appear to travel counterclockwise around the circle (consistent with global motion); when viewed parafoveally, the spots appear to move clockwise (consistent with local motion). When making an eye movement that shifts the right circle from a parafoveal to a foveal location, or vice versa, it appears to suddenly reverse its direction.

Because every element in our display contains both local and global motion, it was natural to ask whether the eccentricity-driven reversal of apparent motion direction happened with only one spot moving around the fixation point. It did not; when all but one of the spots

(Movie: counter-rotating wheels)

MOVIE 1. When fixating at the center of the left wheel, both wheels appear to move in the same direction. But when fixating the center of the right wheel, both wheels appear to move in opposite directions. The appearance of the right wheel's movement reverses depending on the viewing eccentricity.

in the circle were eliminated, the remaining spot appeared to move consistent with its global position shift, regardless of eccentricity. This suggests that interactions between the elements play an important role in causing the local motion to dominate when viewing the stimulus parafoveally. In other words, under conditions of crowding the local percept dominates.

Crowding is a phenomenon wherein identification or discrimination of an object presented in the visual periphery is impaired by the presence of nearby, but non-overlapping flanking objects. A finding characteristic of crowding is that critical spacing (usually a measure of the distance between target and flanker which achieves a particular elevation of threshold for recognition) scales linearly with retinal eccentricity [Bouma, 1970, Toet and Levi, 1992]. Although most studies of crowding focus on its effect of impairing the recognition of shapes (e.g. letters) in parafoveal vision, it has become apparent that crowding is a more general phenomenon, extending to many different types of visual features (e.g. (van den Berg et al. 2007; for review, see Levi 2008) It is thought that crowding is characteristic of some cortical mechanism that integrates signals from low-level feature detectors, a so-called “integration field” [Pelli et al., 2004]. Because the scaling of critical distance with spacing mirrors the variation of cortical magnification with eccentricity, the integration field is thought to be a process that subsumes a constant distance on the cortical surface [Pelli, 2008].

Pelli et al. [2004] proposed that the crucial diagnostic test for crowding as opposed to masking or other forms of spatial interference is that the critical spacing scales with eccentricity and is relatively unaffected by signal size. Accordingly, we set out to determine which target spacing and motion parameters are necessary to drive the reversal of apparent motion as various eccentricities. In section 3 below, we determine the relationship between critical spacing and target spacing, which satisfies Bouma’s law. In section 4 we show that the critical spacing is unaffected by the presence of an occluder which covers 2/3 of the visible circle, meaning that it is the spacing which is relevant and not the number of visible targets. In section 5 we show that the critical distance and scaling property is robust to the size of the stimuli. We also test its robustness to variations in temporal frequency, step size, and step interval.

While most studies of crowding involve stationary stimuli, motion stimuli add a temporal component. In Experiments 1 through 3 we consistently find that for stimuli near the crowding distance, the trials for which the subject took longer in responding were more likely to correctly reflect the global direction of motion. In Experiment 4 we use an auditory cue to vary the subjects’ response time to investigate this effect in more detail. Our results reinforce the idea that global motion processing is the result of an integration of the output of low-level feature detectors, and that in fact the process subserving detection of global motion might be identical to the processes underlying object recognition and target selection. We discuss the implications for possible mechanisms of higher order motion perception and speculate on their possible physiological implementations.

## 2. GENERAL METHODS

**2.1. Subjects.** Five subjects took part in this series of experiments. The subjects and the experiments they took part in are listed in Table 1. Subject P.M. is an author. Subject S.K. was made aware of the purpose of the experiments only after completing section 3. Subjects S.M., D.T., and G.B. were paid and were naive to the purpose of the experiments.

**2.2. Equipment.** Stimuli were presented on a flat CRT video monitor (ViewSonic PF790; 800 × 600 pixels; display area 341 × 256 mm; 120Hz refresh rate) Experiments were programmed in MATLAB using the Psychtoolbox [Brainard, 1997] and Eyelink toolbox extensions [Cornelissen et al., 2002], along with custom OpenGL code. All stimuli were presented on a 50% gray background whose luminance was  $33.10\text{cd}/\text{m}^2$ . The display had a black level of  $0.10\text{cd}/\text{m}^2$  and a white of  $66.05\text{cd}/\text{m}^2$  measured against the gray background.

Subjects sat behind a blackout curtain so that ambient illumination was mostly due to the monitor and viewed the screen binocularly using a chin and forehead rest with the eyes 60 cm from the screen. Eye position was monitored using a video-based eye tracker (EyeLink 1000; SR Research) using a sample rate of 250 Hz. Eye movements were recorded but are not reported in this paper. Subjects gave responses by turning a knob (PowerMate; Griffin Technologies) with their preferred hand.

It would be cute to  
pull this out from  
data but I have more  
pressing concerns.

**2.3. Stimuli.** Example stimuli are shown in 2 and are illustrated in an  $(x, t)$  plot in Figure 1c ( $x$  here being a slice around a circle centered on the fixation point and passing through the center of each motion element.) The stimuli consisted of discrete local motion elements presented at regular temporal and spatial intervals as in apparent motion. Each local motion element had a luminance profile along the circle given by a Cauchy filter function [Klein and Levi, 1985] with peak spatial frequency  $f$ . The luminance profile shifts phase with a constant temporal frequency  $\omega$  and is temporally modulated by a Gaussian envelope with standard deviation  $d/2$ . In the radial direction, each local motion element had a Gaussian envelope with standard deviation  $w/2$ . The equation describing the luminance profile of a patch as a function of position and time is then:

$$C(x, y, t) = \cos^n(\tan^{-1}(fx/n))\cos(n \cdot \tan^{-1}(fx/n) + \omega t)e^{-(t/2d)^2 - (y/2d)^2}$$

with the direction of motion along  $x$ . The spatial bandwidth parameter  $n$  was set to 4 for all stimuli.

Is the equation totally necessary? Is the use of  $x, y$  to describe a circular stimulus confusing here?

(Movie: three example stimuli)

MOVIE 2. Three example stimuli. Subjects viewed stimuli such as these and were asked to judge the direction of movement of the elements (here clockwise in all cases) regardless of the direction of local motion (clockwise, neutral counterphase, or counterclockwise)

In each trial, a number of identical elements were arranged in a circle around the fixation point, each oriented with the direction of motion tangential to the circle. Each element was presented repeatedly at intervals of  $\Delta t$ , each successive appearance displaced a fixed distance  $\Delta x$  around the circle. The examples in 2 have the following settings, the same as used in Experiment 1 :  $\Delta t = 100$  ms,  $\omega = 10$  cyc/s,  $d = 0.033$  s, and if  $\phi$  denotes eccentricity, then  $f = 8.9\text{cyc}/\phi$ ,  $\Delta x = 0.05 \cdot \phi$ , and  $w = 0.066 \cdot \phi$ . The contrast of the local motion elements was 100% for trials using counterphase stimuli, and 70.7% for other trials (so as to keep the motion-energy of the display constant.)

For Experiments 1, 2, and 3, subjects were required to respond within a fixed temporal window. If the latency from motion onset to response was outside the window, the fixation point changed color (red for late responses, blue for early responses) for 1 second as feedback and the trial was reshuffled into the stimulus set to be repeated later in the session.

Subjects performed the task in sessions of at most 1 hour, divided into 4 or 5 blocks of 150 to 200 trials each, and were prompted to take a break between blocks. Subjects could also rest at any point by simply delaying fixation. At the beginning of each block, the eye tracking system was automatically recalibrated by asking the subject to make saccades to a sequence of targets at randomly chosen locations on the screen.

For all experiments reported here, three stimulus types were used with equal probability. In one third of trials the direction of local motion was congruent with that of global motion. In the second third, the direction of local motion was opposite to the direction of global motion. In the remaining trials, elements with counterphase local motion were used. Counterphase elements were constructed by superposing two local motion elements with equal and opposite directions of local motion.; i.e. the counterphase stimuli have the same spatial and temporal frequency content as the congruent and incongruent elements, but their motion energy is equivocal between opposite directions. The second stimulus in 2 shows counterphase local motion.

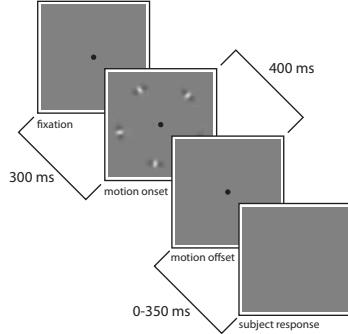
The sigmoids need to be refit with a constant slope (since that is how we will approach the variations and occlusions data; QUEST data doesn't well support calculating slope and it adds noise to the PSE calculation. Also, look at different choices of scaling in element spacing to see what is the best fit? Maybe plot on log but fit linearly? Does that scale?

### 3. EXPERIMENT 1. SCALING OF CRITICAL SPACING WITH ECCENTRICITY.

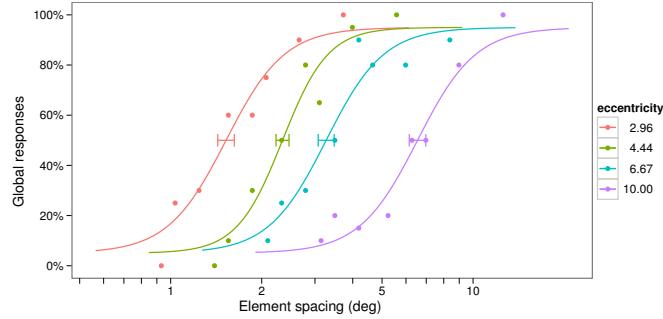
**3.1. Methods.** We presented stimuli at eccentricities of 10, 6.67, 4.44, and 2.96 degrees, using the parameters described above for Movie 1 but scaling all spatial parameters ( $\Delta x$ ,  $w$ ,  $1/f$ ) of the elements along with the eccentricity; i.e. at eccentricity of 6.67 degree,  $\Delta x$  and  $w$  decreased to 2/3 the value used at 10 degrees and  $f$  increased to 3/2 its value. The global apparent motion was shown over 4 stations at intervals of  $\Delta t = 100$  ms.

At each eccentricity, we varied the number of elements in the circle (and consequently the inter-element spacing) in the circle using the method of constant stimuli, using values chosen for each subject based on preliminary sessions.

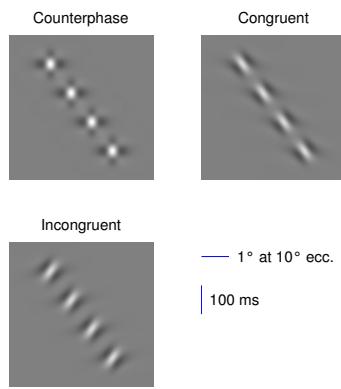
#### 3.2. Results.



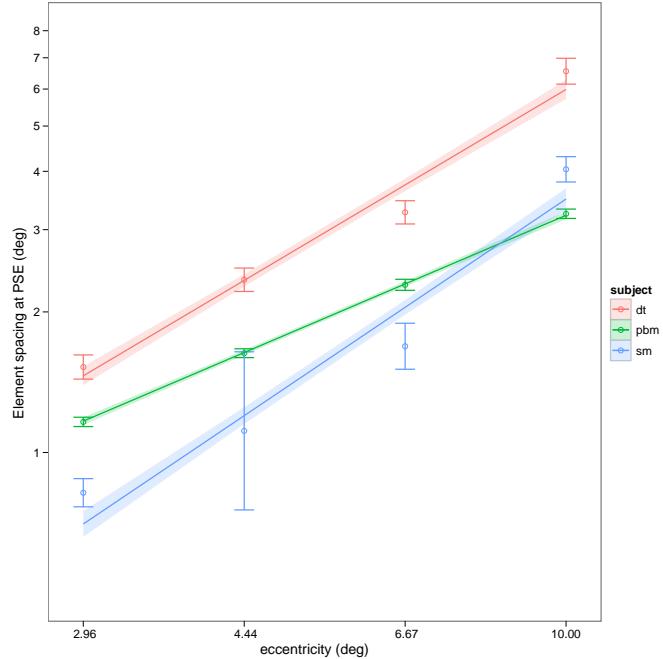
(A) Task illustration. Subjects first fixate and after a brief delay a motion stimulus of constant duration appears. Subjects judge the apparent direction of motion and respond by turning a knob before the time window has expired. Subject receives feedback about whether their response falls inside the time window.



(B) Example data. The responses of observer DT to incongruent stimuli are plotted as a function of between-element spacing, for four values of eccentricity. The values plotted are the proportion of responses that agree with the global motion direction. Curved lines are fit to the data by a cumulative logistic with a constant guess rate. The point of subjective equality (PSE) is indicated on each fit.



(C) Example stimuli in space-time form, where time progresses down along the vertical axis. Stimuli were ‘congruent’, ‘counterphase’ or incongruent, based on whether the global direction of motion agreed with the local. Counterphase stimuli are a superposition of congruent and incongruent stimuli.



(D) Points of subjective equality (target separation subtending the global direction of motion serving 50% response probability) for each eccentricity for all subjects. Intervals show standard errors. Lines show a power-law fit between eccentricity and critical target separation of congruent and incongruent stimuli. Shaded region shows standard error of the power-law fit.

FIGURE 1

Using the third of trials where global motion opposed local, we obtained a psychometric function relating the target spacing to the probability that the stimulus is seen to rotate in the direction of the global motion. The data were fit to a cumulative logistic function using a maximum likelihood estimator. We found the spacing where the logistic curve intersected 50%.

We fit the subject's responses at each eccentricity to a logistic function, as illustrated for subject D.T. by the curves in Figure 1b. A separate curve was fit for each eccentricity, with a guessing rate that was fit for each subject [Wichmann and Hill, 2001]. From these fits we estimated the point at which subject responses would be equally split between local and global directions of motion. This point of subjective equivalence (PSE) is indicated by the horizontal error bars in Figure 1b, and are plotted using vertical error bars in Figure 1d for all subjects. This spacing at the PSE appears to scale with the eccentricity of the stimulus. We made another fit to a model where the size of the PSE was proportional to the stimulus eccentricity; this model fit is shown as lines and shaded regions in Figure 1d. When compared to estimates taken at each individual eccentricity, we saw XXXX significant differences at YYYY conditions.

The scalar dependence on critical separation is broadly similar to the phenomenon of crowding, in which recognition or discrimination of a target object is impaired by the presence of flanking objects. It is also suggestive of a cortical mechanism. There are several areas of cortex that are organized into retinotopic maps. The foveated scaling of space within these maps has the property that network interactions that span a constant distance in cortex, including V1, will correspond to interactions in visual space whose distance approximately scales with retinal eccentricity.

#### 4. EXPERIMENT 2. OCCLUSION.

The results of Experiment 1 suggest that the perceived direction of incongruent motion stimuli follows the global translation when stimuli are widely spaced, but follows the local motion direction when stimuli are packed more closely together. The between-element spacing at which the percept changes from global-dominated to local-dominated appears to be roughly proportional to the stimulus eccentricity. As a result the stimulus at PSE has approximately the same number of elements, independent of eccentricity. This leaves open an alternate explanation: observers' correct perception of global motion may be limited by the number of simultaneously visible elements rather than their closeness. Such might be the case if subjects were determined the direction by individuating and tracking distinct elements in the display. Humans appear to have a limited capacity for individuating and tracking multiple targets within a scene [Pylyshyn and Storm, 1988], which might be overwhelmed by displays with large numbers of targets.

The data from Experiment 1 do not distinguish target number from spacing; there is not a consistent trend for the number of targets at PSE to change as a function of eccentricity. If target tracking capacity may be different for different portions of the visual field, as for

actually do this

I'm not happy about calling this 'subjective equivalence' because I'm not sure what the stimuli are 'equivalent' to, they're not metamers. Perhaps a point of equivocation?

OK, so what's the appropriate test? Fit a model plus one data point, at each data point, and see if the added coefficient was a significant change? (ONLY X conditions, significant marked with a star; were these differences explainable?)

does this need a cite?  
"Who defined crowding" is a hard thing to cite.

cite?

I can cite papers for MT and V4 that make the claim that those maps are just simple linear scalings of the V1 map.

example it appears to differ between the lower and upper visual field [He et al., 1996], then Experiment 1 cannot even in principle differentiate an explanation based on target number from one based on density. To distinguish the effect of target number from that of density, we repeated the measurement of critical spacing with and without an occluder that obscured most of the targets, reducing their number without changing their spacing.

In retrospect, seeing where this analysis is taking me, a much better approach would be to (1) not bother testing at different eccentricities and (2) test varying sizes of the (visible) window!

#### 4.1. Methods.

Trial structure and timing were as in Experiment 1, but with the addition of an occluder that was present or absent on either side of the screen. The occluder was a ‘C’ shape with an opening subtending 120° on either the left or right of the screen and covering all eccentricities that a target might appear at. The stimuli are shown schematically in the legend of Figure 2a, but for the data reported here, the occluder was not visible (i.e. it had the same luminance as the background.)

We decided to test stimuli in the left and right hemifields over separate sessions. If subjects used attentional resources to individuate and track the motion elements, then their spatial allocation of attention may impact performance. Mixing left-visible and right-visible stimuli in the same trial blocks would require subjects to alternately allocate attention to left and right sides, perhaps unnecessarily limiting their performance; preliminary data suggested that a fully interleaved design inflated the critical spacing measurement for occluded conditions. Therefore we measured critical spacing for right and left hemifields in separate sessions. In each session, half of trials contained an occluder and half did not. Note that spatial allocation of attention may also impact subjects’ performance in the case that individuation of elements turns out not to be necessary to perform their task.

(and this inflation was partially recovered by using a visible occluder that cued the side that targets were to appear on. – supplementary figure?)

Instead of the method of constant stimuli we used the QUEST procedure [Watson and Pelli, 1983] to select the target density for each trial. A separate QUEST sequence was used to seek the PSE at each eccentricity and for both fully visible and occluded stimuli (so 8 interleaved QUEST sequences for each session) with trials from all sequences being randomly interleaved. As before, there were three trial types, with local motion congruent with global motion, incongruent with global motion, or in counterphase. The online QUEST estimates were used to select target spacing for all trials, but only incongruent trials with responses given during the response window were used to update the online estimates.

In modeling subjects’ responses we employed a generalized linear regression using binomial errors and a modified logistic link function with variable upper and lower asymptotes, corresponding to a guess rate and a lapse rate [Wichmann and Hill, 2001]; these rates were allowed to vary between subjects and were iteratively fit to maximize likelihood. Comparisons between nested models were done by likelihood ratio test using a  $\chi^2$  distribution with the appropriate number of degrees of freedom depending on the models, as well as by comparisons of Akaike’s information criterion (AIC) between models. Before comparing

I have not actually done this yet; currently it's just a fixed 5% guess rate.

models, guess and lapse rates were first fit to maximize the joint likelihood of both models. For statistics such as the PSE that are not linear functions of the regression coefficients, standard errors and comparisons were computed by parametric bootstrap.

**4.2. Results.** The measured PSEs are shown in Figure 2a. The online estimates from QUEST are disregarded and the PSEs are calculated by fitting a logistic response function with slope, guess and lapse rates varying per subject, as in section 3.

If the perceived motion direction were driven by element number rather than element spacing, we might first expect the measured critical spacing to decrease when most targets are occluded. For Subject DT the smallest spacing at PSE were observed in the session where the unoccluded window was in the left hemifield. In the left hemifield spacing at PSE did not reliably differ between occluded and unoccluded conditions (parametric bootstrap,  $p = .16$ ) In the right hemifield, the spacing at PSE increased for the occluded conditions ( $p = .001$ ). [Author PBM did not show a significant difference in critical spacing between occluded and unoccluded stimuli]

On the other hand, GB showed a significant decrease in critical spacing for occluded stimuli shown on both sides of the screen ( $p < .001$ , left;  $p < .001$ , right.) Over all eccentricities and both hemifields, the average PSE for occluded conditions was  $76\% \pm 14\%$  that of unoccluded conditions. This reduction is considerable, but nonetheless there are more elements visible in one hemifield during unoccluded stimuli at PSE than there are in the occluded condition at PSE. In other words the reduction in PSE is not great enough to be accounted for solely by target number.

Nonetheless, in this first measurement, introducing an occluder has mixed effects on the resulting PSE measurement. This may reflect different strategies taken by each subject, or other differences between subjects. Unpacking the components of the display in the occluded condition may reveal an underlying explanation or the different effects, as well as shedding further light on the question of whether it is target number or spacing that is responsible or the reversed direction of perceived motion at high densities.

### 4.3. Is density a better predictor than number?

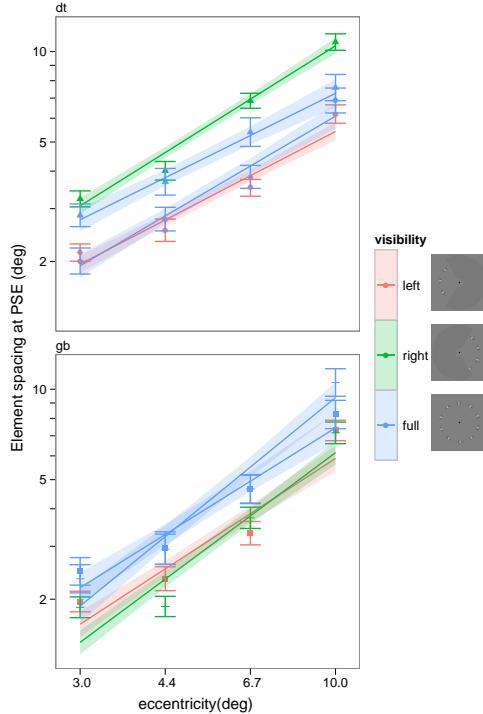
It is the case that multiple values of target density can result in the same number of elements being present. Contrariwise, the same target density in two different trials can result in different numbers of targets being visible, depending on the randomly selected initial position of the targets (Figure 2b). This fact may give us leverage to answer the question of whether it is element density or number that determines the perceived direction of motion, since multiple values of target density have been tested with the same number of visible targets and vice versa.

ref?

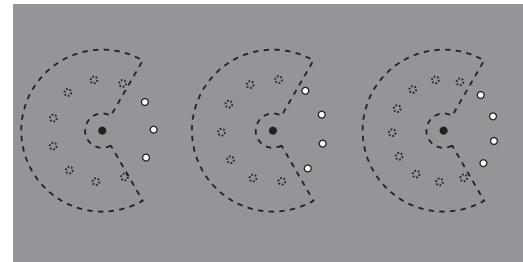
this is apparently the  
case for most data  
but I need to re-run  
using the final set of  
stimulus parameters

this standard error  
includes both the es-  
timation uncertainty  
and the scatter be-  
tween different condi-  
tions, all in one pool.  
Is that confusing?

This section wanted to start with addressing the density vs. num-  
ber fidence ellipsoids it seems that the approach isn't helpful after all.  
you have to dive into appearances and disappearances to untangle.



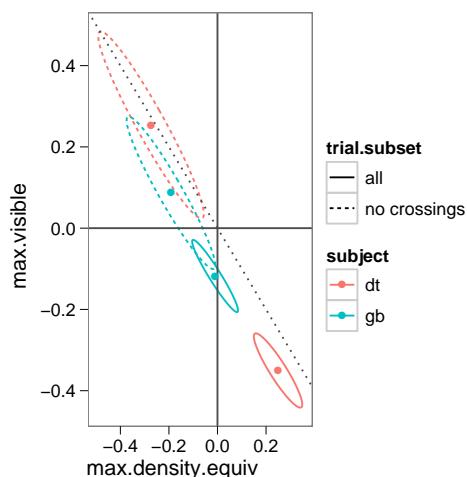
(A) PSE values, measured as in Figure 1. visible in the non-occluded window. Middle: Colors indicate visibility condition (left- With a different starting position, four targets visible, right-visible, fully-visible). Shape gets are visible with the same target density. of data points indicates sessions data were Right: Again four targets are visible, but at taken on; each measurement in a partially a higher density (13 targets in the full circ- occluded condition is matched with a fully- cle). Thus target density and the number of visible condition taken during the same ses- visible targets are not strictly dependent on sion. each other.



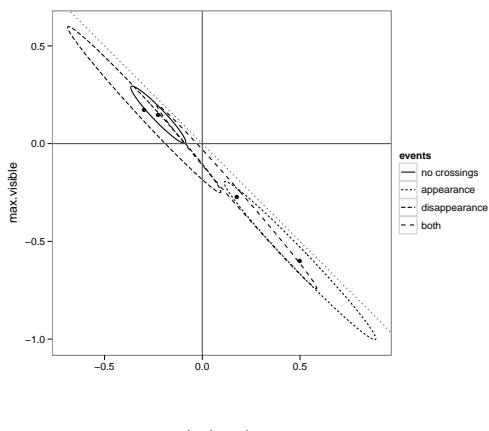
(B) Adding an occluder helps to disentangle target number from target density. Left: out of 11 targets spaced evenly around a circle surrounding the fixation point, two are

visible in the non-occluded window. Middle:

Colors indicate visibility condition (left- With a different starting position, four targets



(c) Confidence ellipse of the coefficients resulting from logistic regression analysis for each subject. The horizontal axis shows the coefficient  $\beta_d$  for target density while the vertical shows the coefficient of  $\beta_n$  for target number. The graph is scaled to equalize the importance of each variable. Solid ellipses show the results for when all occluded tri- als are considered; dashed ellipses consider lipes for the different trial groups according only trials where an element did not cross to whether trials contained appearances or



(d) With subject data pooled, confidence el- lipses are shown for the different trial groups according to whether trials contained appearan-

In order to even discuss this simplified first stab at the question I need to explain why use  $n_{max}$  versus some other count of visible elements. I don't really see a way around it.

As expanded on below, because elements in the display move behind an occluder, the number of targets visible can change during the stimulus. So the notion of ‘number of elements visible’ for each trial can have different definitions. We considered several measures of element number: the number of targets visible at the onset of motion stimulus ( $n_{on}$ ); the number of targets visible at the offset of the stimulus ( $n_{off}$ ); the maximum number of targets visible at any point during the stimulus ( $n_{max}$ ); the minimum number of targets visible at any point during the stimulus ( $n_{min}$ ); and the average number of targets visible integrated over the duration of the stimulus ( $n_{mean}$ ). Of these variables, by far the best fit to subjects’ responses was obtained using  $n_{max}$ ; the difference in AIC score from the next best contender was 26. Therefore we use  $n_{max}$  for this and subsequent analyses in this section.

Note that if you add appearances and disappearances to the model, all these variables span the same space, so once there’s appearances and disappearances, there’s no real distinction between any of these variables other than which ones soak up how much variance from certain appearances and disappearances. So am I shooting myself in the foot here, by picking a “number” alternative that already has some of the effect of appearances and disappearances built in?

We reconstructed the positions of the elements for each trial and calculated  $n_{max}$ . Then we regressed subjects’ responses against two variables,  $n_{max}$  and the corresponding element density  $d_{max}$ . The equivalent element density was defined as the expected maximum number of elements visible in a trial with a given density, taking the expectation over the randomized starting position of the elements. Therefore,  $d_{max}$  is purely a function of the spacing between targets, and trial-by-trial differences between  $d_{max}$  and  $n_{max}$  are due to the randomized starting position of the elements. Regression coefficients of the two variables should thus be directly comparable.

The estimated values of the two regression coefficients are plotted in Figure 2c for each subject, along with standard errors drawn as solid ellipses. As expected, since  $n_{max}$  and  $d_{max}$  are largely correlated, the ellipses are elongated along an axis with slope  $-1$ , indicating that the two coefficients trade off against each other. The estimated coefficients lie below the diagonal line with slope  $-1$ , confirming that either added targets, or added density, tended to produce fewer correct answers. However we do not appear to have a consistent answer as to whether density or number of targets produces a better prediction. For instance, in the case of subject DT, the estimated regression coefficients would have an increased number of targets strongly associated with incorrect answers, while increased density strongly promotes correct answers, a situation which is somewhat nonsensical, unless there is a confounding property of the stimulus that correlates with changes of density and number. Therefore we turn our attention to other visual cues that are present in the stimulus.

“Additionally the residual deviance is probably larger than expected due to overdispersion, indicating that other factors influencing subject’s responses have not been accounted for.” – quantify?

**4.4. Appearances and disappearances as confounding cues.** Introducing an occluder unavoidably introduces ancillary visual cues that can give away the global direction of motion even if crowding obscures the direction of motion of flanked elements. Appearance and disappearance events may provide one cue. As illustrated in (Figure 3a), if elements cross the boundary of the trial window, their appearances or disappearances can provide a significant clue to the direction of global motion, since it is global and not local motion direction that determines where appearances and disappearances happen.

Subjects may also be able to exploit the fact that the elements nearest the boundary of the occluder are only flanked on one side; the effect of crowding is much reduced on targets with only one flanker versus targets with flankers on either side [Bouma, 1970], so that even if the elements in the midst of the window cannot be spatially distinguished, the ‘endpoints’ of a moving mass of elements might still be successfully tracked thereby giving away the global motion direction. In this scenario, appearance and disappearance events would be deleterious, as they cause the ‘endpoints’ to shift opposite the direction of global motion.

In Figure 2c the dashed ellipses show the results of the regression when only trials that do not contain a boundary crossing (appearance or disappearance) are considered. The ellipses are larger due to the reduced sample size, but the data shows that for both subjects, the reversal of apparent motion direction is more likely to be driven by target density than by number. Moreover, when subjects’ behavior is stratified by the presence of boundary crossings in the visual stimulus, the differences between subjects’ behavior appears to be more consistent with each other, up to a change in intercept. Therefore data from both subjects are pooled in Figure 2d. We see that in aggregate, the data from trials containing no crossings is more consistent with apparent motion reversal being driven by target density than by target number; when all trials are considered, it is the appearances in particular that disguise this fact. We proceeded to try to account for the effect of appearances and disappearances in the regression model.

R tells me that there is still a difference between subjects on appearances, but not on the other three groups

**4.5. The effect of appearances and disappearances.** Of the trials with an occluder reported on in Figure 2a, 21% of trials contained an element appearance, 21% of trials contained an element disappearance, 10% of trials contained both, and 68% contained neither. No trials contained more than one appearance or disappearance event.

This is where the model exploration used to start, before the business with the ellipses. Move the data in this para up or eliminate it?

We first added the predictor  $n_{max}$  which was the maximum number of targets visible during the stimulus presentation.  $n_{max}$  was a significant predictor of subjects’ responses (Likelihood ratio test,  $p < 10^{-4}$ ). Furthermore,  $n_{max}$  was a significant predictor among each subset of trials delineated by subject and occlusion condition; (dtleft,  $p < 10^{-4}$ ; dtright,  $p = .013$ ; gbleft,  $p = .0013$ ; gbright,  $p < 10^{-4}$ ). However adding interaction terms for subject and hemifield did not improve the fit, so we did not block  $n_{max}$  by those factors.

We then considered the effect of targets appearing from behind the occluder, adding the regressor  $n_{app}$ , the number of appearance events during each trial. The coefficient for  $n_{app}$

was negative ( $\beta_{n_{app}} = -1.5 \pm .21$ ), indicating that appearances increased the rate at which the subjects answered incorrectly (i.e. in agreement with the local direction of motion.) Furthermore, there was a significant interaction between number of appearances and subject identity ( $p = .0019$ ), indicating that appearance events affected different subjects' responses differently. The effect of appearances was considerable compared to the effect of adding targets to the display; for subject GB, the coefficient associated with  $n_{app}$  was  $.66 \pm .56$  times the coefficient of  $n_{max}$ ; for subject DT the ratio was  $1.5 \pm .63$ . In other words shifting the initial target position so that an appearance occurs midway through the trial had more effect than adding two or three targets to the unoccluded portion of the display. Differences in the effect of appearances may be due to differing strategies employed by the subjects, and may help to explain the differing effects of occlusion on the PSE values measured for Figure 2a.

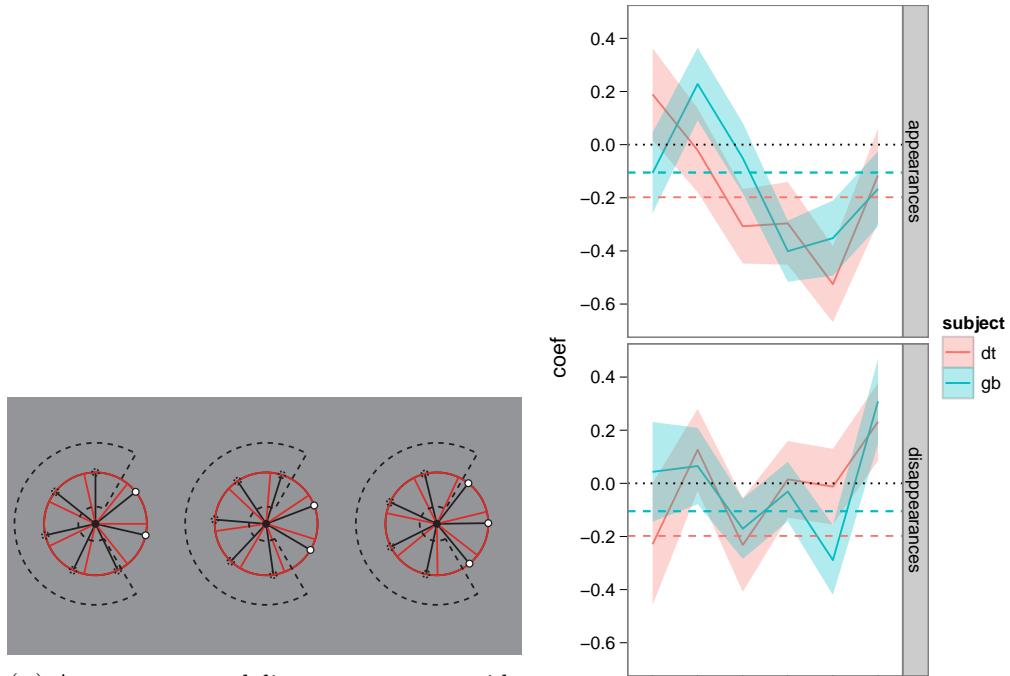
As for disappearances, these also had an effect, although less strongly than appearances. Interestingly, the pattern of effects for disappearances was nearly the inverse of that for appearances: a disappearance generally increased the likelihood of subjects' answering correctly. Moreover, the improvement was significant for subject GB ( $B_{dis,GB} = .2 \pm .28$ ,  $p = .48$ ) but did not reach significance for subject DT ( $B_{dis,GB} = -0.32 \pm .35$ ,  $p = .35$ ). Overall, including disappearances significantly improved the model fit ( $p = .0098$ ) and the fit was not further improved by stratifying by subject ( $p = .22$ ).

I start to think that the small number of trials with both an appearance and a disappearance are so wacky that they throw a lot of fits off, and we might benefit from leaving them out altogether. But what's a good way to quantify that?

Another thing to think about: I described how appearances could plausibly have either direction of effect (inference of global direction vs. movement of the endpoints.) If in fact both effects are present, that may complicate things and a symptom would be excessive residual deviance in the appearance trials.

Since appearances have such a marked effect on subjects' responses, we next looked at the timing of appearances within each stimulus presentation. Appearances that occur immediately following stimulus onset or immediately preceding stimulus offset ought not to have much effect on subjects' responses, since the visual stimulus is virtually identical to one that contains no appearance events. On the other hand, appearances that occur during the middle of the stimulus presentation should have a greater effect. To examine the effect of occlusion events as a function of time we stratified the appearance times into equally sized bins. The effect of appearances and disappearances as a function of time is shown in Figure 3b; for both subjects, the effect of appearances starts near zero and grows significantly negative (associated with responses in the direction of local motion) in the second half of the stimulus presentation. In contrast, the effect of disappearances is smaller, not significantly deviating from zero in these time bins except possibly at the end of the trial.

finally, adding density to the equation; show that it obliterates the need for  $n_{max}$ .



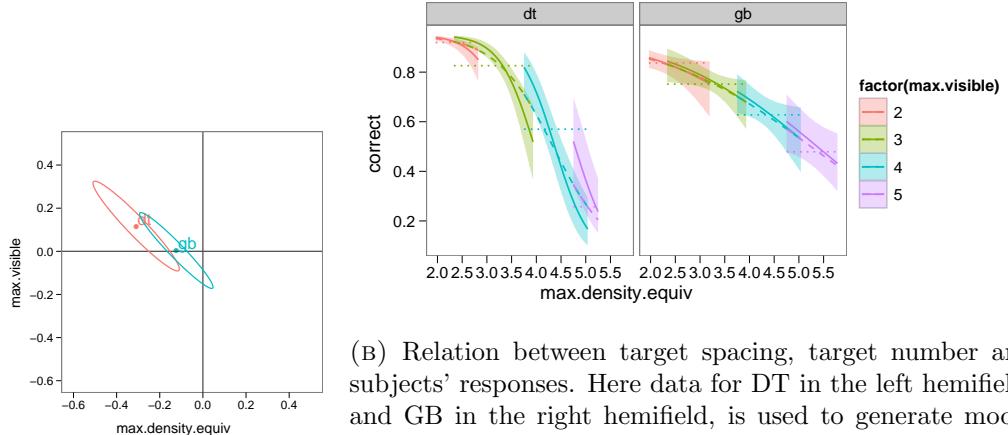
(A) Appearances and disappearances provide additional cues for the direction of motion. The initial position of the targets is shown and the path of their travel is shown as red arrows. Left: a target appears from behind the occluder while the stimulus is visible. Middle: neither an appearance or disappearance occurs during the limited duration of the stimulus. Right: a target disappears behind the occluder while the stimulus is visible. The target density is the same for all three stimuli.

(B) The effect of appearances on subjects' response rate as a function of time. The horizontal axis divides the stimulus duration into bins. The vertical axis plots the change, in log odds of subjects' rate of answering according to global motion direction, as compared to stimuli that do not contain an appearance. Bars indicate standard error of the estimated effect.

FIGURE 3. Effect of appearance and disappearance events in partially occluded stimuli.

I got stuck on how-to-get-to-this-figure for a bit, since . But then I discovered that a stepwise regression starting with a kitchen-sink of the various factors ends up spitting out a model that drops target-number and keeps target-density as the most significant explanatory variable. Stepwise regression doesn't know anything about the problem domain and proceeds dumbly by significance value ignoring effect size, so I'd like to not use it and have a more traditional discussion, but since a dumb algorithm gets to the right place, that lets me synthesize the key figure.

Overall, then, in our model which attempts to account for the side effects of occlusion, it appears that element density explains the apparent reversal of motion direction better than



(A) Confidence ellipses as in spacing while holding the number of targets visible in the Figure 2c for the occlusion window. Different colors correspond to different numbers model, showing the values of of targets visible in the window, while the equivalent target both  $\beta_d$  for target density spacing is plotted along the horizontal axis. Dotted lines while the vertical shows the indicate predictions made using only element number, and coefficient of  $\beta_n$  for target dashed lines show a fit that uses only target spacing. Solid number. Both subjects' fits lines and shaded confidence regions are fit allowing a mixture models' results are more consistent with target density than with target number. The mixture models' results are more compatible with those of the target spacing than those of target number.

FIGURE 4

element number. We visualize this in two ways. One is to replot the ellipsoids of Figure 2c; in Figure 4a, these ellipsoids now draw on data from all occluder trials and account for occlusion side effects, and are more consistent with a main effect of target density than one of number (graphically speaking they place more density on the x-axis than on the y-axis). Another is to draw the predicted psychometric functions resulting from the model, showing how these functions vary in terms of both target number and target density, and what the uncertainty in the prediction is. We do this in Figure 4b. Here we compare model fits obtained under three conditions: one in which the subjects' response depends on number (dotted lines), and a model allowing a mixture of the mixture. The same set of interaction variables are used in all fits. In Figure 4 we only show the data from each subject's better hemifield (left for DT, right for GB.) We observe in that in Figure 4a, for both subjects, the measured coefficients are more consistent with the illusion being driven by element density rather than number. In Figure 4b we see that when allowing a mixture of number and density to be used in fitting the psychometric function, the result (solid lines) follows more closely a function based only on density (dashed lines) than one based only

on element number (dotted lines). Thus we assert that the apparent reversal of motion direction is driven by element density.

and as a conclusion, now I really ought to update Figure 2a for the more detailed model.

## 5. EXPERIMENT 3. ROBUSTNESS OF SCALING TO STIMULUS PROPERTIES.

We have characterized the interference between moving elements in the display leading to reversal of perceived motion in terms of their spatial separation, and have observed that the critical spacing at which opposing local motion interferes with global motion appears to scale proportionately to eccentricity. However, the size and speed of the elements in our display were also scaled proportionate to their eccentricity. This leaves open an alternate explanation in which the motion reversal is dependent on the distance between the objects *relative* to their size of the objects. Under this hypothesis, the critical spacing might be better described in units of a multiple of the size of the envelope, or the period of the grating, rather than the separation between centers. For example, contour integration among Gabor patches has been characterized as a function of the ratio of inter-element spacing to envelope size [Beaudot and Mullen, 2003] In this section, we vary the size the motion elements, as well as their local and global speeds, and measure what effects these parameters have on the critical distance.

A second motivation is to solidify the connection between our motion stimulus and the phenomenon of crowding. As we have seen the scaling of critical distance with spacing in our display is similar to that observed for crowding. A noted property of crowding in parafoveal vision is that the range of spatial interaction is not strongly dependent on the size of the target or the flankers [Levi et al., 2002, Tripathy and Cavanagh, 2002]; this has been proposed as a diagnostic test for crowding as opposed to other kinds of suppressive spatial interaction [Pelli et al., 2004]. If motion integration is subject to the same effects of crowding as other forms of spatial feature integration, it may reflect a common computational principle or even a shared neural mechanism underlying such processes.

Last sentence probably belongs in Discussion.

**5.1. Methods.** Subjects repeated the direction discrimination task, with the same basic structure and equal proportion of congruent, incongruent and counterphase stimuli as in previous sections. As in ?? we used the QUEST procedure to select the target density (expressed as the number of targets in the full circle) for each trial. Within each session we varied one of the stimulus parameters across three values, those being 66%, 100% and 150% the values used in section 3. Stimulus sizes were otherwise scaled proportionate to eccentricity, as in section 3. QUEST trials were interleaved across three stimulus values and four eccentricity values, so that 12 QUEST sequences were interleaved in each session. Data from separate days were pooled.

*I'm adding this. It will put the motion energy analysis on more solid ground.*

5.1.1. *Stimulus variations and scaling.* We tested three variations of the stimulus: temporal frequency, spatial period, and spatial step size (Figure 5). Since we leave the spatial bandwidth parameter constant, lengthening the spatial period also lengthens the envelope of the individual motion elements, and since the temporal frequency is held constant, lengthening the period also increases the speed of the local motion component of the stimulus. The second manipulation is a change in temporal frequency of the stimulus. This has the effect of changing the speed of the local motion component of the stimulus, without changing its size or spatial frequency. The third manipulation is a change in the step distance between each appearance of the local element. This changes the speed of the global translation, leaving local speed unchanged.

A fourth manipulation is to directly manipulate the balance of local motion energy content. We do this by superposing congruent and incongruent local motion in the same stimulus, in the same way that we constructed our counterphase stimulus by superposing congruent and incongruent local motion stimuli.

*I suppose it's not "bandwidth" but Q-factor of the filter. The filter width contains more oscillations at higher frequencies is the main point.*

5.1.2. *Motion energy analysis.* For each unique configuration of the stimulus values, up to a rotation, we calculated the luminance values along the circle passing through the element centers, sampled at the monitor frame rate. To this data we applied a motion-energy model, similar to Adelson and Bergen [1985], using space-time separable analysis filters. The spatial analysis filters were intended to approximate the bandpass properties of direction selective channels in human vision; to that extent we used a set of Cauchy filters [Klein and Levi, 1985] with the center frequencies matched to those of the stimuli. The spatial bandwidth of the filters was dependent on the center frequency and the eccentricity, consistent with the bandpass properties inferred from human psychophysics; in particular, at a given eccentricity, the bandwidths of the filters decrease with increasing spatial frequency [Anderson and Burr, 1987, Banks et al., 1991, Anderson et al., 1991]. The bandwidths chosen for the analysis filters were thus a function of both eccentricity and center spatial frequency; we selected the spatial extent of each analysis filter by interpolating the measurements of Banks et al. [1991]. The parameters used at each eccentricity are given in ???. On the other hand, the temporal component of each filter was identical at all cases, because temporal sensitivity does not appear to vary with spatial frequency or eccentricity [Virsu et al., 1982, Wright and Johnston, 1983]. The temporal component of each filter was the same as used in Kiani et al. [2008].

## 5.2. Results.

### 5.2.1. Element size.

*Ignore subject GB here. He can barely see the local motion (misses a lot of congruent catch trials at high density.) Looking at his data did convince me to start thinking about the balance of global motion discriminability and local motion energy. But I don't think i can use his data.*

In the first row of Figure 6 we see that the approximately proportional relation of stimulus size to critical distance is maintained at the target size is enlarged or shrunk. There is no significant change in the shape of this slope, however there may be a small enlargement of the critical spacing as the target size is increased. Note that the absolute stimulus size still scales with eccentricity in this experiment; the stimulus size is adjusted relative to

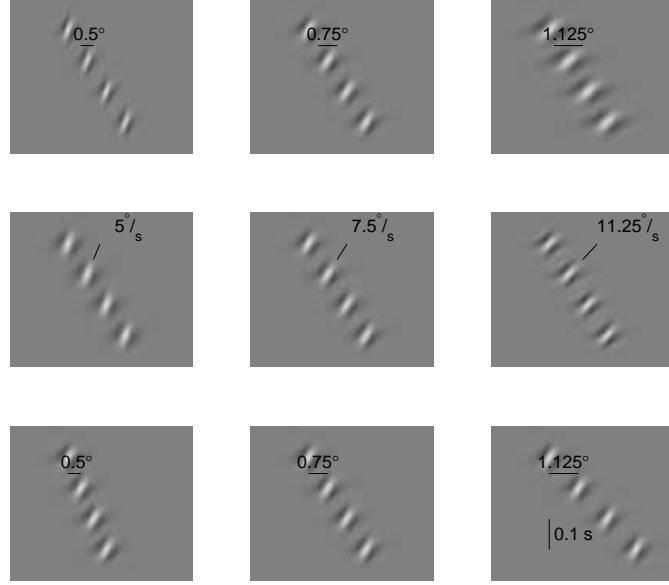


FIGURE 5. Stimulus manipulations.  $(x, t)$  diagrams of the stimuli show the manipulations undertaken. Top row: the element size and spatial period increase from left to right. Middle row: The local speed and temporal frequency increase from left to right. Bottom row: The step size increases from left to right. Stimuli in the middle column are identical. Note that the manipulations of element size and step size produce stimuli that are spatially compressed or elongated versions of each other.

the baseline of sizes used in section 3. In the first row of section 7a we plot the critical distance as a function of the absolute stimulus size. This allows us to see that while holding element size constant over a range of eccentricities, that critical spacing still scales with eccentricity. The small shift of critical distance with target size is on the same order as the element size itself. This result is consistent with crowding because our critical distance has been measured from element center to element center. For large flankers, it is typically the edge-to-edge distance between target and flanker at critical spacing that is invariant to target size [Levi et al., 2002].

**5.2.2. Local speed.** We then measured how critical distance changed as a function of the local speed, which was varied by changing the temporal frequency of phase rotation in each local motion element. The second rows of Figure 6 shows that the scalar dependence of

Again ignoring GB.  
But it's interesting  
that GB collapses  
into size-determined  
behavior when he  
can't see the local  
motion.

This also may touch  
on the question of the  
slope of the psycho-  
metric function. Do  
larger flankers result  
in different slopes? If  
so, then the partic-  
ular threshold level  
I choose for "criti-  
cal distance" affects  
the description of the  
data.

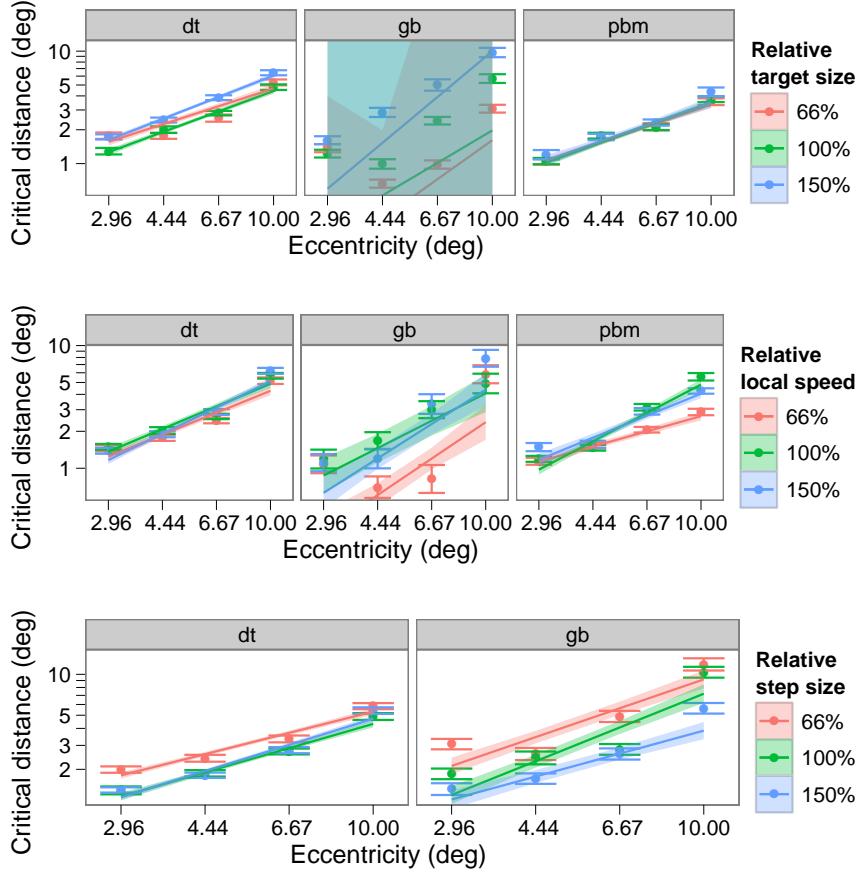


FIGURE 6. Effect of stimulus on critical distance. Each plot shows critical distance for each subject, as a stimulus parameter is varied. Error bars show standard error of the measured critical distance at each condition. Lines show best fit to a power law for critical distance as a function of eccentricity; shaded region shows standard error on that fit. Rows correspond to rows in Figure 5; columns correspond to subjects; colors correspond to stimulus parameters (columns in Figure 5.) Conditions where critical distance varies significantly as stimulus parameters change are labeled with \* ( $p < 0.05$ )

critical distance on eccentricity does not significantly change as local speed changes. This indicates to us that if the

and section 7a show how the measurement of

5.2.3. *Step size.* We then changed the step size of the global translation of each element. We found that the critical separation actually decreased somewhat with increasing step size (section 7a, row 3). At first glance it is difficult to interpret this finding. The result is somewhat contrary to our intuitions; if “misbinding” were taking place between a the present instance of a target and a true future instance of a distractor, a larger step size would be seen to bring the flanker closer to the previous position of the true target and then lead to an increased critical spacing for increased step sizes. However it is not clear that this intuitive temporal structure (a local motion leading to an prediction of future position shift leading to a misbinding) reflects the true computation taking place. Before speculating too deeply on mechanism we must account for the changes in strength of the local motion signal and the global position-change signal. That is the subject of the next section.

this speculation probably belongs in Discussion.

5.2.4. *Strength of local and global motion in various stimuli.* The point can be regarded as the. While changing the flanker distance between widely spaced targets does not affect their local motion energy, crowding has been seen as a kind of equivalent of adding noise about the position of basic features. Thus the ability to detect. We see this in the one third of our trials that use ambiguous “counterphase” motion stimuli that

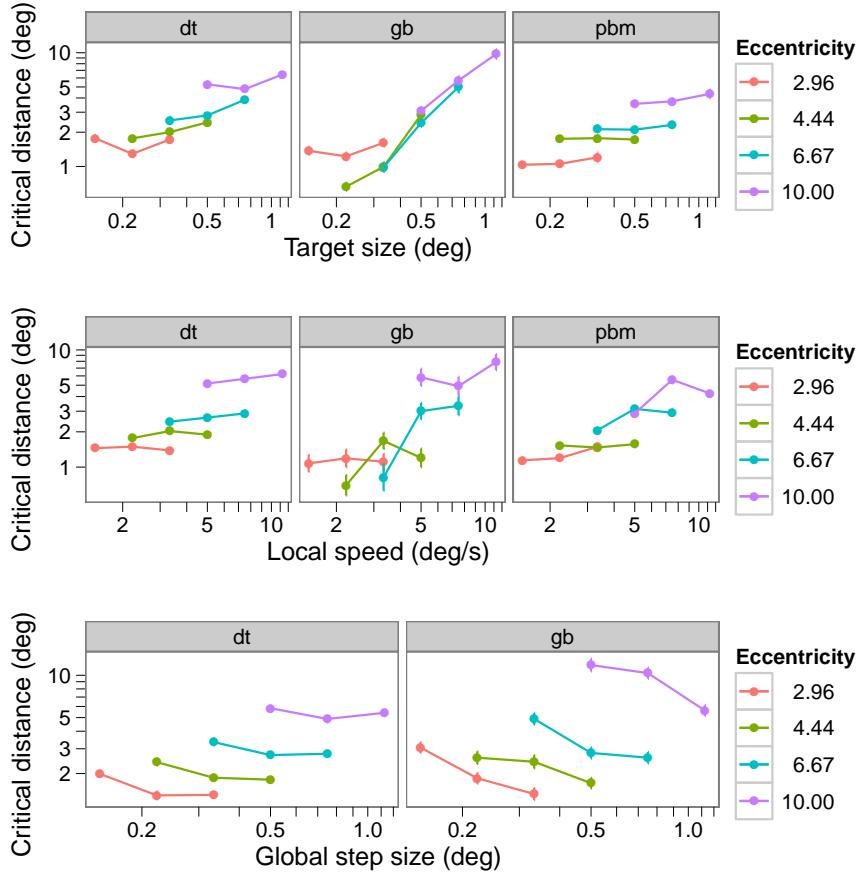
While widely separated flankers do not appreciably affect the local motion energy of the stimuli, the interactions between successive appearances of the local motion elements do have an affect. When we change the size of the local motion elements, or the global step size, we affect the amount of spatial overlap between the successive elements. A local motion detector, such as a complex cell in area V1, whose receptive field is located between the location of two successive presentations of a local element, may be affected by the degree of overlap. It may also be affected by whether the successive elements have their peaks and troughs aligned so that they constructively or descriptively interfere; all elements in our displays were presented in cosine phase, so out manipulations affect the alignment of peaks between successive elements. We performed a motion energy analysis as described in subsubsection 5.1.2.

An experiment titrating the amount of motion energy – putting stimuli on a continuum between “incongruent” and “congruent” by varying the contrast of each – will help me out here as well.

## 6. DISCUSSION

This section is not organized; more of a dumping ground for prompts.

In a field of dots undergoing random Brownian movement, a single dot that changes its position in a consistent direction is more easily detected than can be explained by local motion detectors situated along the object’s trajectory and treated as independent samples [Vergheze et al., 1999]. The enhancement in detectability seems to occur only after 100 ms of movement, possibly due to a process that responds to an initial cue from local motion detectors by reducing the number of detectors monitored to those in the vicinity of the initial motion signal, in particular those in the object’s predicted path [Vergheze



(A) Critical distances from ?? are plotted again, with the horizontal coordinate being the physical size of the parameter being varied as indicated in Figure 5. Rows correspond to rows in Figure 5.

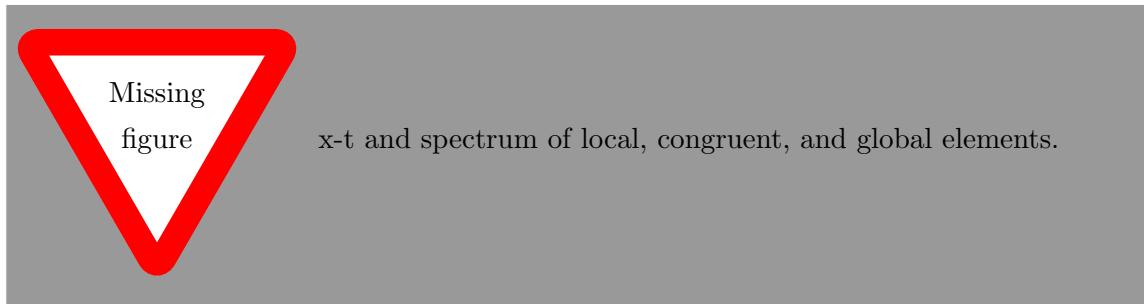
I don't think the error bars ought to be this small; I blame the way that the slopes of the psychometric function have been fixed per-subject.

Missing  
figure

Plot critical distance for SF data as ratio to stimulus size

(B) For the experiment varying target size, critical distance is replotted as a multiple of the target size.

FIGURE 7



(A) Fourier spectra corresponding to local and global motion. Dotted outlines show the full width at half maximum of filters used in our motion energy analysis. Left shows the spectrum of a single element. Middle plot shows the spectrum of congruent motion. Right plot shows the spectrum of incongruent motion. Dotted outlines show the passband (full width at half maximum) of the analysis filters used to extract the motion energy.

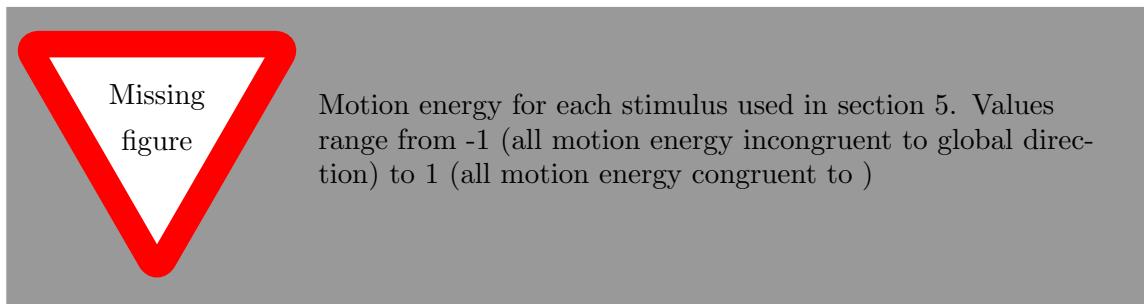


FIGURE 8

and McKee, 2002]. An interaction between local motion and position sensors thus appears necessary to account for performance at motion discrimination.

Discuss crowding as a source of uncertainty about (relative) position

Discuss anatomical implications of eccentricity rule. Many brain areas are foveated. v4 receptive field, v1/lateral connection size, other numerology

why is it so hard to find global motion in the brain? We seem to locate it among a class of phenomena which integrate local visual features into higher level percepts. What about the ventral stream?

Having two classes of stimuli that are well distinguished behaviorally, we would hope to be able to find neurophysiological correlates of both types of motion perception. The underpinnings of local motion are Vaina and Cowey [1996] provides one positive finding, a patient with a unilateral cortical lesion slightly posterior to the hMT+ complex, resulting in a deficit in detection of global but not local motion in the contralateral field. Because

MS: I expected to see a broader conclusion statement about synergy, cooperativity, or interaction... without returning the focus to the specific case involved. If you want to limit, as in some cooperative interaction ... at play... even in simple detection

receptive fields in MT are large, reflecting the integration of many V1 receptive fields spanning a range of spatial positions, it is natural to suppose that a global motion process could be supported by MT. However, attempts to observe MT neurons in the act of responding to global, as opposed to local, motion, have met with little success. For example, when random dot local motion stimuli are presented in a window that moves independently of the dots, MT cells respond primarily to the local motion [Priebe et al., 2001]. Responses in macaque MT and MST to a stimulus opposing local and global motion showed no selectivity of cell responses to global motion direction, even though the stimuli elicited an oculomotor pursuit response in the direction of global motion [Ilg and Churan, 2004]. Livingstone et al. [2001] used sparse noise to map second-order spatiotemporal kernels in MT receptive fields and could not find any spatiotemporal interaction at scales larger than those of V1 receptive fields. Finally, recordings in MT made using stimuli similar to those used in this report find no selectivity in MT cells for global motion [Shadlen et al., 1993, Hedges et al., 2004]. It appears that direction selective responses in MT are, like those of V1, a function of local, and not global, motion.

Verghese and McKee [2002] found that detectability of a single dot with consistent direction among randomly moving backgrounds was enhanced after 100 ms of target motion. In that report they also show (Figure 1 of that paper) that there is some enhancement is preserved even if the target motion is discontinuous; a target that suddenly jumped sideways or backwards in the middle of its motion trajectory still had enhanced detectability relative to the baseline. However, if the size of the jump was too large, the enhancement vanished. In the light of the present study, we suggest that the critical maximum jump size may be the same as the critical spacing of crowding. The enhancement observed in that paper has been interpreted as a process that responds to an initial cue from local motion detectors by reducing the number of detectors monitored to just those in the vicinity of the initial motion signal. It may be that the mechanism that winnows the pool of motion detectors is the same as the integration field of Pelli et al. [2004].

## REFERENCES

- E H Adelson and J R Bergen. Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A*, 2(2):284–299, 1985. ISSN 0740-3232 (Print).
- S J Anderson and D C Burr. Receptive field size of human motion detection units. *Vision Res*, 27(4):621–35, 1987.
- S J Anderson, D C Burr, and M C Morrone. Two-dimensional spatial and spatial-frequency selectivity of motion-sensitive mechanisms in human vision. *J Opt Soc Am A*, 8(8):1340–51, Aug 1991.
- M S Banks, A B Sekuler, and S J Anderson. Peripheral spatial vision: limits imposed by optics, photoreceptors, and receptor pooling. *J Opt Soc Am A*, 8(11):1775–87, Nov 1991.
- William H A Beaudot and Kathy T Mullen. How long range is contour integration in human color vision? *Vis Neurosci*, 20(1):51–64, 2003.

- H Bouma. Interaction effects in parafoveal letter recognition. *Nature*, 226(5241):177–8, Apr 1970.
- DH Brainard. The psychophysics toolbox. *Spatial Vision*, 10(4):433–436, 1997.
- Frans W Cornelissen, Enno M Peters, and John Palmer. The eyelink toolbox: eye tracking with matlab and the psychophysics toolbox. *Behav Res Methods Instrum Comput*, 34(4):613–617, 2002. ISSN 0743-3808 (Print).
- S He, P Cavanagh, and J Intriligator. Attentional resolution and the locus of visual awareness. *Nature*, 383(6598):334–7, Sep 1996. doi: 10.1038/383334a0.
- JH Hedges, A Kohn, SC Saint, NC Rust, MN Shadlen, and JA Movshon. Neurons in macaque mt signal local but not global motion. Society for Neuroscience abstract:Program No. 526.522., 2004.
- Uwe J Ilg and Jan Churan. Motion perception without explicit activity in areas mt and mst. *Journal of Neurophysiology*, 92(3):1512–1523, 2004. ISSN 0022-3077 (Print). doi: 10.1152/jn.01174.2003.
- Roozbeh Kiani, Timothy D Hanks, and Michael N Shadlen. Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J Neurosci*, 28(12):3017–29, Mar 2008. doi: 10.1523/JNEUROSCI.4761-07.2008.
- S A Klein and D M Levi. Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation. *J Opt Soc Am A*, 2(7):1170–1190, 1985. ISSN 0740-3232 (Print).
- Dennis M Levi. Crowding—an essential bottleneck for object recognition: a mini-review. *Vision Res*, 48(5):635–654, Feb 2008. ISSN 0042-6989 (Print). doi: 10.1016/j.visres.2007.12.009.
- Dennis M Levi, Srividhya Hariharan, and Stanley A Klein. Suppressive and facilitatory spatial interactions in peripheral vision: peripheral crowding is neither size invariant nor simple contrast masking. *J Vis*, 2(2):167–177, 2002. ISSN 1534-7362 (Electronic). doi: 10:1167/2.2.3.
- M S Livingstone, C C Pack, and R T Born. Two-dimensional substructure of mt receptive fields. *Neuron*, 30(3):781–93, Jun 2001.
- Z L Lu and G Sperling. The functional architecture of human visual motion perception. *Vision Res*, 35(19):2697–2722, 1995. ISSN 0042-6989 (Print).
- Denis G Pelli. Crowding: a cortical constraint on object recognition. *Curr Opin Neurobiol*, 18(4):445–451, 2008. ISSN 0959-4388 (Print). doi: 10.1016/j.conb.2008.09.008.
- Denis G Pelli, Melanie Palomares, and Najib J Majaj. Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J Vis*, 4(12):1136–69, Dec 2004. doi: 10:1167/4.12.12.
- N J Priebe, M M Churchland, and S G Lisberger. Reconstruction of target speed for the guidance of pursuit eye movements. *Journal of Neuroscience*, 21(9):3196–3206, 2001. ISSN 1529-2401 (Electronic).
- Z W Pylyshyn and R W Storm. Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spat Vis*, 3(3):179–97, 1988.
- Michael N Shadlen, E Zohary, Kenneth H Britten, and William T Newsome. Directional properties of MT neurons examined with motion energy filtered apparent motion stimuli. *Investigataive Ophthalmology and Visual Science Supplement*, 34(4):908, 1993.

- Robert M. Steinman, Zygmunt Pizlo, and Filip J. Pizlo. Phi is not beta, and why Wertheimer's discovery launched the gestalt revolution. *Vision Research*, 40(17):2257–2264, 2000. URL <http://www.sciencedirect.com/science/article/B6T0W-40VT194-4/2/ccd5af03d347e42d9fb50131930ed666>.
- A Toet and D M Levi. The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Res*, 32(7):1349–1357, 1992. ISSN 0042-6989 (Print).
- Srimant P Tripathy and Patrick Cavanagh. The extent of crowding in peripheral vision does not scale with target size. *Vision Res*, 42(20):2357–69, Sep 2002.
- L M Vaina and A Cowey. Impairment of the perception of second order motion but not first order motion in a patient with unilateral focal brain damage. *Proc Biol Sci*, 263 (1374):1225–1232, 1996. ISSN 0962-8452 (Print). doi: 10.1098/rspb.1996.0180.
- Ronald van den Berg, Jos B T M Roerdink, and Frans W Cornelissen. On the generality of crowding: visual crowding in size, saturation, and hue compared to orientation. *J Vis*, 7(2):14.1–11, 2007. ISSN 1534-7362 (Electronic). doi: 10.1167/7.2.14.
- P Verghese, S N Watamaniuk, S P McKee, and N M Grzywacz. Local motion detectors cannot account for the detectability of an extended trajectory in noise. *Vision Research*, 39(1):19–30, 1999. ISSN 0042-6989 (Print).
- Preeti Verghese and Suzanne P. McKee. Predicting future motion. *J. Vis.*, 2(5):413–423, 9 2002. ISSN 1534-7362. URL <http://journalofvision.org/2/5/5/>.
- V Virsu, J Rovamo, P Laurinen, and R Näsänen. Temporal contrast sensitivity and cortical magnification. *Vision Res*, 22(9):1211–7, 1982.
- A B Watson and D G Pelli. Quest: a bayesian adaptive psychometric method. *Percept Psychophys*, 33(2):113–20, Feb 1983.
- F A Wichmann and N J Hill. The psychometric function: I. fitting, sampling, and goodness of fit. *Percept Psychophys*, 63(8):1293–313, Nov 2001.
- M J Wright and A Johnston. Spatiotemporal contrast sensitivity and visual field locus. *Vision Res*, 23(10):983–9, 1983.