Rapid serial visual presentation of motion: Short-term facilitation and long-term suppression

Padma B. Iyer  
School of Medical Sciences, University of Sydney, Australia

Alan W. Freeman  
School of Medical Sciences, University of Sydney, Australia

J. Scott McDonald  
School of Psychology, University of Sydney, Australia

Colin W. G. Clifford  
School of Psychology, University of Sydney, Australia

The visual system can detect coherent motion in the midst of motion noise. This is accomplished with motion-sensitive channels, each of which is tuned to a limited range of motion directions. Our aim was to show how a single channel is affected by motions both within and outside its tuning range. We used a psychophysical reverse-correlation procedure. An array of dots moved coherently with a new, randomly chosen, direction every 14 or 28 ms. Human subjects pressed a key whenever they saw upwards movement. The results were analyzed by finding two motion directions before each key-press: the first preceded the key-press by the reaction time, and the second preceded the first by a variable interval. There were two main findings. First, the subject was significantly more likely to press the key when the vector average of the two motions was in the target direction. This effect was short-lived: it was only seen for inter-stimulus intervals of several tens of milliseconds. Second, motion detection was reduced when the target direction was preceded by a motion of similar direction 100–200 ms earlier. The results support the idea that a motion-sensitive channel sums sub-optimal inputs, and is suppressed by similar motion in the long term.

Keywords: motion-2D, temporal vision, computational modeling


Introduction

The visual system often has the task of detecting one motion direction among others present at the same time and place. How does one perceive a flock of birds flying across a background of drifting clouds? How does one analyze the multitude of motions perceived when moving through a cluttered environment? Our ability to pick a target motion direction out of motion noise is well illustrated by studies of random dot motion in which a small fraction of the dots move in the same direction while all other dots move in random directions. Primate subjects are able to perform this task when as few as 2% of the dots move together (Newcombe & Paré, 1988).

Two themes have emerged from previous study of motion direction discrimination. First, objects moving with similar directions tend to produce a perceiver of motion in their average direction. Williams and Sekuler (1984) used dots that randomly varied their directions over time. When the directions were chosen from a distribution spanning angles less than about 180° the overall motion was perceived to be in the direction of the distribution mean. This was not the case when the direction was widened: local motion then failed to produce a coherent motion perceiver. The cooperative effect of similar motions extends to detection. Simpson and Newman (1998) showed that two successive motions were more easily detected when the motions were in similar directions.

The second theme to emerge from previous work is of suppressive perceptual interactions between opposing motions. Qian, Andersen, and Adelson (1994) used two arrays of dots moving in opposite directions. The dots were placed so that each dot in one array was paired with an opposing dot from the other array. Surprisingly, the percept was not of one array moving transparently over the other, but of flicker. The authors surmised that a motion perceiver was absent because the paired motion signals cancelled each other out, abolishing global motion. When opposing dots were spatially offset by at least 0.2°, however, transparent motion was restored. The authors concluded that unpaired displays send unbalanced directional signals leading to a transparent perceiver. On the other hand, when the dots were paired they cancelled each other’s motion signal, indicating a local suppressive interaction.

While this previous work has demonstrated the existence of cross-motion interactions, it has provided incomplete answers to a major question: what are the time courses of the interactions? Motion processing must have a fast component in order to deal with rapid movement. There are also indications that suppressive interactions are on a slower time scale (Snowden, 1989). We have addressed
this question with an experimental design using a rapid series of stimuli (Busse, Katzner, Tillmann, & Treue, 2008; Neri & Levi, 2008; Potter, 1975; Ringach, 1998; Tadin, Lappin, & Blake, 2006). A field of dots was translated as a group and assigned a random motion direction at a rate of either 36 or 72 Hz. An analysis of motion detection demonstrated not only facilitatory and suppressive interactions between motions, but also their time courses. This work has been presented previously in abstract form (Iyer, Freeman, Clifford, & McDonald, 2009).

Methods

Subjects

Five subjects took part in the study. All were aged between 27 and 38, and four were female. The subjects wore their usual optical correction, if any, and had a visual acuity of better than 6/6 and a stereo-threshold less than 1 min. One of the subjects was also an author (PI). The other subjects were unaware of the aims and results of the study.

Stimuli

Stimuli were presented on a computer monitor. The software used to generate the stimuli, collect subject responses, and analyze results, ran within Matlab (The MathWorks, Inc.) and included functions in the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The monitor was located 1.14 m from the subject, and its screen subtended 13° horizontally by 10° vertically at the eye. The screen had a spatial resolution of 1 pixel/min and a frame rate of 72 Hz. The stimulus background was white \((x = 0.288, y = 0.308)\) with a luminance of 64 cd/m², and the room lights were off. Subjects used a chin- and forehead-rest to reduce head movements. The stimulus was enclosed within a black border, as shown in Figure 1. The inner dimensions of the border were 2.5° × 2.5° and border width was 0.25°. A white dot 0.1° in diameter was placed at the center of the bordered area; both border and dot helped to stabilize fixation. The stimulus comprised 30 black dots, each dot being 0.1° in diameter. The luminance of the border and black dots was 0.9 cd/m² and that of the fixation dot was 115 cd/m².

At the start of an experimental run the black dots were randomly distributed within a 2° × 2° area centered in the bordered area: both the horizontal and vertical location of each dot was randomly selected from a uniform probability density 2° in width. The dots were shifted as a group from one scene to the next. Scene durations were 28 ms (2 video frames) or 14 ms (1 video frame) with scene rates of 36 and 72 Hz, respectively. The distance moved was set so that the velocity of the apparent motion was 3 deg/s. The motion direction on each scene was selected with equal probability from 20 possible directions evenly distributed across the full 360° range. When a dot moved outside the stimulus area it was relocated to the opposite side of that area.

Procedure

Each run lasted 60 s. One motion direction—vertically upwards—was nominated as the target, and subjects

![Figure 1](https://jov.arvojournals.org/pdfaccess.ashx?url=data/journals/jov/933483/)
pressed a key whenever they saw motion in that direction. Data were collected from each subject over 10 1-hour sessions, with rest breaks during each session as needed. At least 200 minutes of data were collected from each subject for runs of the 36 Hz movie, and at least 240 minutes for the 72 Hz movie.

**Data analysis**

Our analysis procedures follow and extend those used by previous studies in which a series of stimuli were presented in rapid succession, and responses were collected during stimulation (Busse et al., 2008; Perge, Borghuis, Bours, Lankheet, & van Wezel, 2005; Ringach, 1998). Analysis commenced by finding the motion directions preceding a key-press: an example is shown in Figure 2A. This graph shows data from a single subject presented with a 36 Hz movie. It was obtained by finding the motion direction 389 ms (14 video frames) before each key-press, and compiling this sample of directions into a frequency histogram. Frequency was converted into probability by dividing by the number of key-presses. Repeating this analysis for a range of times prior to each key-press produced the surface plot in part 2B of the figure. Time is shown on the vertical axis and probability by the gray level. A numerical guide to gray...
level is provided at the right of the graph. Densities obtained about 400 ms prior to a key-press peak at the target motion direction, indicating a correlation between stimulus and response. Densities at shorter and longer times are more or less flat, indicating a lack of correlation. Given the tight link between a target stimulus and the following response, we assume a causal relationship and therefore refer to the interval between stimulus and response as the reaction time. The reaction time for this subject is about 400 ms.

The reaction time can be described more fully by finding those times at which the probability density differs significantly from flatness. A chi-square statistic was therefore calculated for each density using the following formula:

\[
\chi^2 = \sum_{i=1}^{20} \frac{(n_i - m)^2}{m},
\]

where \(n_i\) is the number of key-presses preceded by motion direction \(i\), and \(m\) is the mean number of key-presses (\(\sum n_i/20\)). Figure 2C plots chi-square as a function of time prior to a key-press for all five subjects. Four subjects have a reaction time of around 400 ms and one subject about 550 ms. Although this subject’s reaction time was unusual, her results from the remaining analyses were consistent with those of the other subjects.

Figure 2D shows the same data as in part C of the figure but with an expanded vertical scale. The dashed line shows the 5% significance level for chi-square as obtained from a goodness-of-fit test. Chi-square is clearly significant over a range of times prior to a key-press. We used this observation to collapse the multiple probability densities into a single one: each significant probability density was weighted by its chi-square (less the significance level), and the weighted densities were averaged. It may be noted that there are a number of significant values at very short (about 100 ms) and long (around 800 ms) times prior to a key-press. These contributions were very small compared with those around the reaction time, and had a negligible effect on the weighted average. Figure 3 shows the result, with one line per subject.

For the 36 Hz movie target motions were presented at a mean rate of 1.8 Hz (=36/20) and the mean key-press rate was 0.36 Hz across all subjects. Subjects therefore responded to only about one fifth of targets. For the 72 Hz movie, the rates were 3.6 Hz and 0.38 Hz respectively, meaning that subjects responded to about one tenth of targets.

## Results

### Interactions between consecutive motions

The main aim of this study is to describe how the detectability of one motion is influenced by a preceding motion. To this end, subjects viewed a rapid stream of random motions, as illustrated in Figure 1. In the first experiment, the dots moved as a group 36 times every second. Subjects pressed a key whenever they saw the target motion, which was vertically upwards. These responses were analyzed by finding the motion directions that preceded a key-press. The first-order analysis described in the Methods section shows that the most probable motion preceding a key-press was the target direction, and that the reaction time was 400 ms or more.

The reaction time can be described more fully by finding those times at which the probability density differs significantly from flatness. A chi-square statistic was therefore calculated for each density using the following formula:

\[
\chi^2 = \sum_{i=1}^{20} \frac{(n_i - m)^2}{m},
\]

where \(n_i\) is the number of key-presses preceded by motion direction \(i\), and \(m\) is the mean number of key-presses (\(\sum n_i/20\)). Figure 2C plots chi-square as a function of time prior to a key-press for all five subjects. Four subjects have a reaction time of around 400 ms and one subject about 550 ms. Although this subject’s reaction time was unusual, her results from the remaining analyses were consistent with those of the other subjects.

Figure 2D shows the same data as in part C of the figure but with an expanded vertical scale. The dashed line shows the 5% significance level for chi-square as obtained from a goodness-of-fit test. Chi-square is clearly significant over a range of times prior to a key-press. We used this observation to collapse the multiple probability densities into a single one: each significant probability density was weighted by its chi-square (less the significance level), and the weighted densities were averaged. It may be noted that there are a number of significant values at very short (about 100 ms) and long (around 800 ms) times prior to a key-press. These contributions were very small compared with those around the reaction time, and had a negligible effect on the weighted average. Figure 3 shows the result, with one line per subject.

For the 36 Hz movie target motions were presented at a mean rate of 1.8 Hz (=36/20) and the mean key-press rate was 0.36 Hz across all subjects. Subjects therefore responded to only about one fifth of targets. For the 72 Hz movie, the rates were 3.6 Hz and 0.38 Hz respectively, meaning that subjects responded to about one tenth of targets.

![Figure 3. Probability densities measured at different times prior to a key-press were weighted and summed to obtain the densities shown here. The lines, one for each subject, show that the most probable motion before a key-press is in the target direction.](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933483/ on 11/19/2018)
The marginal density $p_1(d_1)$ in this equation was obtained by summing the observed probabilities across values of $d_2$ and, similarly, $p_2(d_2)$ was obtained by summing probabilities across $d_1$. (Examples of $p_1(d_1)$ and $p_2(d_2)$, obtained more directly, are shown in Figure 3.) The result of applying Equation 2 is shown in Figure 4B. If $d_1$ and $d_2$ interact in producing a key-press, the independence model $p_{ind}$ will differ from the observations $p_{obs}$. The last step in the analysis therefore subtracted one from the other:

$$p_{interaction} = p_{obs} - p_{ind};$$

and is shown in part C of the figure. All of these calculations were performed for a range of times prior to each key-press. The resulting interaction plots were weighted by their chi-squares (for direction $d_1$, as described in the Methods) and averaged. The plots in Figure 4 and in the following figures were all calculated with this weighted averaging procedure.

The interaction plot is not identically zero, indicating that the two stimuli interact in producing a key-press. Bright areas indicate that the combination of stimuli is more probable than if they acted independently, and are therefore labeled Facilitation. Similarly, dark areas are labeled Suppression. To test for statistical significance, the interaction plot was tested with a two-way analysis of variance. The two factors were $d_1$ and $d_2$, each with 20 levels, and subjects provided the five replicates. The interaction term in this test, $d_1 \times d_2$, was highly significant ($F(361, 1200) = 3.54, p < 0.001$). The most interesting feature of the interaction plot is that the area of facilitation lies on the negative diagonal. This means that, for example, a motion 36° anticlockwise from the target combines with a following motion 36° clockwise from the target to make a key-press more likely. Put otherwise, a target motion is more likely to be seen when the vector sum of consecutive motions is in the target direction.

The interval between the two stimuli analyzed in Figure 4 was 28 ms. The perception of brief motions requires rapid neural processing and we wondered, therefore, whether interactions between successive motions could be found for even briefer intervals. To this end we ran a second experiment, this time with a 72 Hz movie (1 video frame per scene). One subject was not available for this experiment, and will therefore not be shown in subsequent figures. Two interaction plots from this experiment are shown on the right of Figure 5. The upper one represents the case in which the two analyzed motions were consecutive and have an inter-stimulus interval of 14 ms. For the lower plot the two analyzed motions were separated by one intervening motion, and therefore have an inter-stimulus interval of 28 ms. The interaction plot obtained with the slower movie, shown in Figure 4C, is reproduced on the left side of Figure 5 for comparison. Data from the faster movie indicate facilitation on the negative diagonal for an inter-stimulus interval of 14 ms, but not for 28 ms.
These data suggest that the vector summation effect is brief. At the end of the Results section we present a further analysis designed to measure the duration of this effect.

Modeling short-term interactions

What is the source of the vector summation seen in the interaction plots? Single neurons have relatively broad tuning for motion direction (Albright, 1984; Britten & Newsome, 1998; Lagae, Raiguel, & Orban, 1993; Maunsell & Van Essen, 1983): perhaps two motions bracketing the target direction both stimulate a channel tuned for the target direction, producing sub-threshold summation. This idea is illustrated in Figure 6A. The diagram shows tuning curves for three motion-sensitive channels and the dashed lines indicate two motion directions bracketing the target direction. Assume that the bracketing motions are delivered in close succession, and that the channels simply add their responses to those stimuli. The central channel yields the greatest summed response. If the channel tuned to the target direction is assumed to trigger a key-press when its response is greater than that of all other channels, key-presses will result when the vector sum of motions is in the target direction.

We tested this idea by simulating the responses of such a model, which is fully described in Appendix A. An interaction plot for the model is shown in Figure 6B. The simulation used a 72 Hz movie as stimulus and the analysis assumed that stimulus $d_1$ immediately followed $d_2$. The result can therefore be compared with the upper right plot in Figure 5. Facilitatory areas for the model lie on the negative diagonal, lending support to the idea that they are due to summation in a broadly tuned detector. The Discussion expands on these ideas. The model’s interaction plot contains areas of suppression as well as of facilitation. These do not arise from inhibition, as the model contains no inhibitory mechanisms. Rather, they result from the analysis method. The probabilities in each of the observations (Figure 4A) and independence model (Figure 4B) plots sum to unity (the certain event). The values in the interaction plot (Figure 4C) therefore sum to zero, and the presence of facilitation in one part of the plot leads to suppression elsewhere. The most useful aspect of an interaction plot, therefore, is not so much the existence of facilitatory and suppressive areas, but their relative locations within the plot.

Suppressive interactions between consecutive motions

Another interesting feature of the interaction plots in Figure 5 is the consistent presence of a suppressive area at...
the middle of the upper and lower boundaries. The dashed semicircles provide examples: the suppression at the center of these semicircles proved to be significant in a one-sided t-test ($p = 0.047$). This indicates that preceding a target motion with a motion in the opposite direction makes a key-press less likely. This finding fits well with previous observations that motion is less visible when presented along with another motion in the opposite direction (Iyer & Freeman, 2009; Lindsey & Todd, 1998; Mather & Moulden, 1983).

**Time course of cross-motion interactions**

The right side of Figure 5 shows that the pattern of interactions between responses to two motions depends heavily on the interval between them. To explore this observation further, we plot interactions over a much larger range of inter-stimulus intervals in Figure 7. Data from the 36 Hz stimulus movie are shown on the left, 72 Hz on the right. The direction of one motion is shown on the horizontal axis, the direction of the preceding motion on the vertical axis, and the interval between the two motions is shown to the right of each row. The uppermost three plots are reproduced from Figure 5. The remaining plots, for longer inter-stimulus intervals, show a pattern very different from the short-term effect. In particular, a dark area appears at the center of the plot. The mean plot over inter-stimulus intervals of 83–278 ms, presented at the bottom of the figure, shows this central area of suppression clearly. The central suppressive area indicates that the detectability of a target motion is reduced by presenting a motion in the same direction hundreds of milliseconds previously.

To more precisely depict the time course of this suppression, we quantified the interactions at the center of the plot by summing values within the dashed circle shown at the bottom of the figure. The radius of the circle was that which minimized the sum of interaction values within it: smaller circles did not include all suppressive interactions, and larger circles encroached on facilitatory interactions surrounding the central region. This sum of interactions is shown as a function of inter-stimulus interval in Figure 8. Data for the 36 Hz movie are shown on the left, and for the 72 Hz movie on the right. The time course is similar in the two cases: two target motions in close succession facilitate a key-press whereas an inter-stimulus interval of hundreds of milliseconds suppresses a response.

In the interests of separating the facilitation and suppression time courses, we note that there are two clearly distinguishable patterns in the interaction plots of Figure 7. One is the pattern, $b$, observed at the briefest inter-stimulus intervals (28 ms for the 36 Hz movie, and 14 ms for the 72 Hz movie). This pattern is a function of motion directions $d_1$ and $d_2$ and is therefore denoted $b(d_1, d_2)$. The second pattern, $l(d_1, d_2)$, is seen in the long-term and can be characterized by the mean plots shown at the bottom of the figure. This second pattern can be calculated by noting that it is symmetric about the horizontal midline (the line for which $d_2$ is equal to the target direction). Accordingly, the long-term pattern was set equal to the mean over all inter-stimulus intervals, $t$, of the symmetric component of the interaction plot, $i(t, d_1, d_2)$:

$$l(d_1, d_2) = \text{mean over } t(0.5 \times (i(t, d_1, d_2) + i(t, d_1, -d_2))).$$

We assumed that any interaction plot could be modeled as a weighted sum of these two patterns:

$$\text{Interaction}(t, d_1, d_2) = f(t)b(d_1, d_2) + s(t)l(d_1, d_2) + \varepsilon,$$

(5)
where \( f(t) \) and \( s(t) \) are the time courses of the two patterns, and \( e \) is error not accounted for by the model. This model was fitted to the empirical interaction plots by least-squares regression, and the resulting time courses are shown in Figure 9 with \( f(t) \) and \( s(t) \) labeled Facilitation and Suppression, respectively. Data for the 36 and 72 Hz movies are shown on the left and right, respectively, time courses for individual subjects are shown in the upper row, and the decompositions of the mean across subjects in the lower row. F-testing showed that all regressions were significant at the 5% level; 95% confidence intervals are provided. The results show that the facilitatory pattern is limited to inter-stimulus intervals of less than about 100 ms, whereas the suppression between similar

Figure 7. Interaction plots are shown for the 36 Hz movie on the left and the 72 Hz movie on the right. The time interval between the two interacting stimuli increases towards the bottom of the figure and is shown on the right. Data are means over four subjects. Gray levels are scaled on each plot so that the deepest suppression is represented by black and the greatest facilitation by white. The interaction plots at the longest inter-stimulus intervals differ markedly from those at short intervals: there is a suppressive area at the middle of the plot, indicating that a key-press is less likely when a target stimulus is preceded by another target several hundred milliseconds before.
The previous figure shows that two stimuli in the target direction have very different interactions for short and long inter-stimulus intervals. To quantify this change over time, interaction values were summed within the dashed circle shown at the bottom of that figure. The present figure shows the sum versus inter-stimulus interval for the 36 Hz and 72 Hz movies on the left and right, respectively. There is one line for each subject. The time courses were smoothed with a five-point moving average. The plots show facilitatory interactions only for short inter-stimulus intervals and suppressive interactions thereafter.

The curves were calculated by assuming that the interaction plots were linear combinations of the facilitation pattern shown at the top of Figure 7 and another pattern representing long-term suppression. The suppression pattern was calculated by finding the interaction component that was symmetrical about the horizontal midline, and by taking the mean of this component across all inter-stimulus intervals. Data for the 36 and 72 Hz movies are shown on the left and right, respectively. Data for individual subjects are shown at top and the decomposition of the mean across subjects is shown at bottom. All curves were smoothed with a five-point moving average. The error bars are 95% confidence intervals. Facilitation is short-lived relative to suppression.
motions peaks between 100 and 200 ms, and lasts for at least 500 ms.

What is the source of this long-lasting suppression? One possibility is motor: subjects are unable to press the key twice in a short period. The period of interest is the reaction time plus the duration of long-term suppression, a total of about 1 s. To test for motor effects we repeated the analysis illustrated in Figure 9, excluding those key-presses that followed the previous one by less than 1 s. The result was negligibly different from that in Figure 9, making it unlikely that motor effects contribute to the long-term suppression. A stronger possibility for the source of this suppression is a perceptual process. In particular, the suppression time course is similar to that of another phenomenon seen during rapid serial visual presentation, namely the attentional blink (Raymond, Shapiro, & Arnell, 1992). This similarity is taken up in the Discussion.

Discussion

A number of previous studies have investigated the detection of a motion signal when two or more motion directions are present. Earlier work used simultaneously presented motion directions to look at the influence of one motion on the detection of a motion in another direction (Curran & Braddick, 2000; Fredericksen, Verstraten, & van de Grind, 1994; Lindsey & Todd, 1998; Marshak & Sekuler, 1979; Mather & Moulden, 1983; Qian, Andersen, and Adelson, 1994; Rauber & Treue, 1999; Watamaniuk, Sekuler, & Williams, 1989; Williams & Sekuler, 1984). More recent work has used a rapid series of motion stimuli to reveal the time course with which a target motion is detected; subject responses were collected at the end of discrete trials (Neri & Levi, 2008; Tadin et al., 2006) or during the stimulus (Busse et al., 2008). Periods in which the probability of motion detection is reduced can be due to either suppression or reduced excitation, and it is clearly useful to distinguish between these two possibilities. Busse et al. took a step in this direction by using a second-order analysis: target detection was measured as a function of two preceding motions, rather than just one. We have taken a further step by subtracting out independent effects of the two preceding motions. In the process we have gone beyond previous work by demonstrating how the interactions of two motions influence motion detection.

Our analysis identified three periods in the motion processing time course. Initially there is fast cross-direction facilitation, lasting several tens of milliseconds, that promotes perception of a vector sum across motions. Within a similar time frame there is also a suppressive interaction between the processing of opponent motions. The third component is also suppressive but is between motions in the same direction and takes more than 100 ms to peak. We address each of these phenomena in turn.

Short-term facilitation

The first striking feature of the data revealed by our second-order analysis is that the probability of subjects reporting an upward motion is greater than expected when sequential motion vectors sum to vertical upward motion. Previous psychophysical experiments have also suggested some kind of vector summation or averaging. Williams and Sekuler (1984) moved dots in a random direction each frame. When the range of directions was less than about 180° subjects perceived a global motion in the average direction of the dots. Likewise, when Watamaniuk et al. (1989) assigned the path directions of random dots according to a Gaussian distribution, the subjects saw global motion in the approximate direction of the distribution mean. Qian, Andersen, and Adelson (1994) used pairs of moving dots with a direction difference of only 45°, and found that subjects perceived motion along the axis of average motion. Curran and Braddick (2000) repeated this, but also measured motion averaging of paired dots 120° apart. Our contribution is to demonstrate that the vector averaging of motions occurs over several tens of milliseconds.

Putting an exact figure on the duration of the vector averaging effect is difficult for the following reason. We have assessed the strength of this effect by presenting two motions separated by a variable number of intervening motions. As suggested by one of the reviewers of this paper, a lack of summation between the responses of the two motions could occur because of masking from the intervening motions. We have no easy way of measuring any masking effects and, as a result, cannot provide precise timing for the facilitation time course.

We propose a simple model to explain the observed vector summation. It is based on three propositions: 1. subjects have a bank of directionally selective channels; 2. the channels have broad direction tuning but a short integration period; 3. for a subject to respond, the channel most responsive to the target direction must be activated more than the other channels. The output of the model is shown in Figure 6, and clearly predicts the short-term facilitation well. To understand how, first consider the simplest case. When an upward signal is fed into such a system, the channel tuned to this stimulus is activated more than the other channels, prompting the subject to respond. In the case of a single motion not in the target direction, the channel tuned to the target will be activated less than other channels and the subject does not respond. However, when two motions bracket the target direction in close sequence, the successive activations of the channel tuned to the target sum to provide activation greater than that of other channels. The subject then responds to a vector average in the target direction.

Where might these directionally selective channels be found in the brain? MT is thought to be a good candidate for motion processing, because monkeys in which it has been damaged show impairments in motion processing (Newsome & Paré, 1988), and electrical stimulation of MT
biases motion perception (Salzman, Britten, & Newsome, 1990). Physiological studies give strong support to the theory of a bank of channels with broad direction tuning in MT (Albright, 1984; Britten & Newsome, 1998; Maunsell & Van Essen, 1983; Snowden, Treue, & Andersen, 1992). Perge et al. (2005) studied the responses of neurons in macaque MT to a series of motions presented at 75 or 120 Hz. Each motion direction was randomly selected from eight directions spread evenly across the 360° range. A comparison of our Figure 5 with their Figure 8A shows similarities. In particular, Perge et al. found that successive motions bracketing a neuron’s preferred direction interacted to make a following action potential more likely. They also found that this facilitation declined markedly when the two motions were not successive.

**Short-term suppression**

A second feature demonstrated by our analysis is that subjects are less likely to report the upward motion when a downward motion immediately precedes it. The downward motion signal appears to inhibit processing of the upward motion. This is consistent with previous findings: Qian, Andersen, and Adelson (1994) argued that the collapse of motion transparency with closely paired dots moving in opposite directions is a result of local inhibition between motion-sensitive mechanisms. Snowden (1989) also reported reduced detection of apparent motion in a dot field by an orthogonal moving field. Mutual inhibition between mechanisms that process direction can also explain the phenomenon of motion repulsion (Marshak & Sekuler, 1979; Rauber & Treue, 1999), in which two fields of translating dots appear to have a larger difference in directions than is physically the case.

Neurophysiological studies also demonstrate phenomena that are compatible with our finding of opponent inhibition between motion mechanisms. Snowden et al. (1992) demonstrated strong inhibition of MT neuronal responses to the preferred direction of motion when the neurons were simultaneously presented with opponent motion. Their data suggest suppression is strongest in the direction opposite to the preferred direction, as we have found with our data. Qian and Andersen (1994) provided further evidence for the same finding. Perge et al. (2005), however, found evidence for response facilitation when a preferred stimulus was preceded by one moving in the opposite direction; the reason for this inconsistency between results is not clear.

To explain their data, Qian, Andersen, and Adelson (1994) proposed a model in which subunits of MT neurons mutually inhibit one another. Simoncelli and Heeger (1998) formulated a more elaborate model of MT neurons with the intention of accommodating a wider range of experimental results. In their model pattern-sensitive neuronal responses not only underwent subtractive adjustment from non-optimal motion directions, but were also subject to divisive inhibition from pooled responses of neurons sensitive to all directions. Both these models, however, lack explicit predictions for the time course of the inhibition. Again, our data suggest new limits on the temporal dynamics of motion processing: Figure 7 shows that such suppressive mechanisms are complete within 56 ms.

**Long-term suppression**

In a final analysis we investigated the interactions in the processing of motion signals over an extended period. This analysis clearly shows a progression from the early cross-motion interactions discussed above, into suppression of preceding same-motion signals. This new pattern emerges after approximately 100 ms and is largely complete by 500 ms (see Figure 9). Snowden (1989) reported that suppression of apparent motion detection occurs on a timescale of several hundred milliseconds. Neri and Levi (2008) found that motion detection was reduced for a period of 100–200 ms after the onset of a rapid series of motion stimuli, and attributed this reduction to delayed self-normalization of motion signals.

One of the most notable aspects of the long-term suppression is that it follows a time course similar to that of other rapid serial visual presentation phenomena, specifically repetition blindness (Kanwisher, 1987) and the attentional blink (Raymond et al., 1992), both of which are held to result from higher-level processing. Kanwisher (1987) suggested that repetition blindness is a failure to recognize individual targets as separate episodes. The attentional blink is thought to be a result of interference with a following target by its predecessor, although there is no comprehensive account (Dux & Marois, 2009). Until recently repetition blindness and the attentional blink were thought to be limited to complex stimuli such as words (Kanwisher, 1987), letters (Raymond et al., 1992), and concepts (Dux & Coltheart, 2005). We have shown recently, however, that there is a failure to detect a grating when it is preceded by another grating of the same orientation, and the time course of this loss peaks at 100–200 ms (Wong, Roeber, & Freeman, 2010). That result, along with the present one, suggests that the origins of the failure to detect repeated stimuli may be in primary visual cortex, where orientation and direction selectivity arise. The low level loss could then propagate, and possibly amplify, as it passes to those higher cortical levels where words, letters, and semantics are processed.

In conclusion we used randomly translated dot fields in a rapid serial visual presentation task to investigate temporal integration of motion signals. Analysis of our data provides evidence for very fast short-term vector summation of motion signals and inhibition of opponent motions. Furthermore, we find evidence of a longer-term...
suppressive interaction between same-direction motions, which we suggest is the result of both low- and mid-level processing.

Appendix A

Description of the model

A key-press is more likely when the vector sum of prior stimuli is in the target direction. We developed a model to better understand this phenomenon. The model consists of an array of motion-sensitive channels, $i = 1, 2, \ldots, m$, tuned to motion directions, $c_i$, that are evenly distributed across the full 360° of motion directions (see Figure 6A for an illustration). The sensitivity of channel $i$ to motion direction $d$ has amplitude

$$a_i(d) = \exp\left(-\frac{(d - c_i)^2}{(2\sigma_d^2)}\right),$$

where $\sigma_d$ is the standard deviation of its tuning curve. In keeping with previous work (Fredericksen et al., 1994; Simpson & Newman, 1998), each channel is assumed to be a low-pass temporal filter

$$\frac{dr_i(t)}{dt} = s(t) - r_i(t),$$

where $r_i(t)$ is the channel’s response as a function of time $t$, $\tau$ is its time constant, and $s(t)$ is the motion stimulus. The stimulus to channel $i$ is a sequence of motion impulses with direction $d_j$, $j = 1, 2, \ldots, n$, at times $t_j$:

$$s(t) = a_i(d_1)\delta(t - t_1) + a_i(d_2)\delta(t - t_2) + \ldots$$

$$\quad + \exp\left(-\frac{(d_1 - c_i)^2}{(2\sigma_d^2)}\right)\delta(t - t_1)$$

$$\quad + \exp\left(-\frac{(d_2 - c_i)^2}{(2\sigma_d^2)}\right)\delta(t - t_2) + \ldots$$

where $\delta(t)$ is a (Dirac) delta function at time $t$. The time course therefore consists of a step increase at each stimulus, followed by an exponential decay. The size of the step declines as the stimulus direction shifts from the direction to which the channel is tuned.

The time course of the model was calculated by numerical integration of Equation A2 (using Matlab’s ordinary differential equation solver ode45). Simulation time was 200 minutes. The response in the channel tuned to the target direction was monitored, and a key-press was triggered each time this response exceeded that of all other channels. A variable delay was added to the time of the key-press to simulate the variability of reaction times shown in Figure 2. This delay was a random sample from a Gaussian probability density with standard deviation $\sigma_r$.

The analysis of key-press times was the same as that used for the empirical data.

The model parameters

$$m = 20,$$

$$\sigma_d = 54^\circ,$$

$$\tau = 30 \text{ ms},$$

$$\sigma_r = 60 \text{ ms},$$

were set as follows. The number of channels, $m$, is not critical provided it is no less than the number of motion directions, 20. The tuning curve bandwidth, $\sigma_d$, was taken from Britten and Newsome (1998) who measured the bandwidth in a population of MT cells (their Figure 3, stimulus coherence of 100%, bandwidth divided by $\sqrt{2}$ to convert it to a standard deviation). The time constant, $\tau$, was set so that the facilitation was relatively small at an inter-stimulus interval of 28 ms, consistent with the data in Figure 5. The standard deviation, $\sigma_r$, was set equal to that found empirically in Figure 2.

Acknowledgments

This work was supported by a University of Sydney R & D Grant to AF, and an Australian Research Council Discovery Project and Australian Research Fellowship to CC.

Commercial relationships: none.
Corresponding author: Alan W. Freeman.
Email: Alan.Freeman@sydney.edu.au.
Address: P.O. Box 170, Lidcombe, NSW 1825, Australia.

References


