THE ROYAL SOCIETY

doi 10.1098/rspb.2000.1406

Matched filtering in motion detection and discrimination

William A. Simpson* **and Velitchko Manahilov**

Vision Sciences Department, Glasgow Caledonian University, Cowcaddens Road, Glasgow G4 0BA, UK

When humans detect and discriminate visual motion, some neural mechanism extracts the motion information that is embedded in the noisy spatio-temporal stimulus. We show that an ideal mechanism in a motion discrimination experiment cross-correlates the received waveform with the signals to be discriminated. If the human visual system uses such a cross-correlator mechanism, discrimination performance should depend on the cross-correlation between the two signals. Manipulations of the signals' cross-correlation using differences in the speed and phase of moving gratings produced the predicted changes in the performance of human observers. The cross-correlator's motion performance improves linearly as contrast increases and human performance is similar. The ideal cross-correlator can be implemented by passing the stimulus through linear spatio-temporal filters matched to the signals. We propose that directionally selective simple cells in the striate cortex serve as matched filters during motion detection and discrimination.

Keywords: visual motion perception; motion detection; motion discrimination; ideal observer

1. INTRODUCTION

Human observers can readily detect and discriminate the motion of visual patterns. Any effort to understand the neural machinery underlying visual motion perception must start with a model of how the required information is extracted from the spatio-temporal luminance distribution falling on the retina. Previous motion models (Adelson & Bergen 1985; Watson & Ahumada 1985) were constructed with the sole constraint that their output can be used for detecting and discriminating motion. Here we will remove a degree of arbitrariness by deriving an ideal motion detector or discriminator. The ideal observer extracts all the information from the noisy received stimulus and achieves the highest possible performance. After deriving the model, we perform three psychophysical experiments that test whether human motion performance is based on a near-optimal neural mechanism.

2. IDEAL OBSERVER ANALYSIS

In our experiments the observer is presented with one of two discretely sampled spatio-temporal waveforms (moving gratings) in normal white noise and has to decide which was presented. The received waveform is

$$r_{xyt} = \begin{cases} s_{0xyt} + n_{xyt} & \text{or} \\ s_{1xyt} + n_{xyt}, \end{cases} \tag{1}$$

where

$$s_{0xyt} = c\cos(2\pi fx + v_0 t), \tag{2}$$

$$s_{1xyt} = c\cos(2\pi fx + v_1 t), \tag{3}$$

and c is the contrast, f is the spatial frequency and v_0 and v_1 are the drift speeds in radians s^{-1}; the normal white noise n_{xyt} has mean zero and standard deviation σ. There are three standard motion tasks used in psychophysics. In *motion detection* the observer discriminates a stationary (v_0 is zero) from a moving ($v_1$is not equal to zero) grating. In *speed discrimination* the observer discriminates gratings moving at different speeds (v_0 and v_1 are unequal and have the same sign). In *direction discrimination* the observer discriminates gratings moving at the same speed but in opposite directions ($v_0 = -v_1$). Space–time plots of s_0 and s_1 for each task are shown in figure 1.

The ideal observer cross-correlates the received waveform with each of the expected signals (Whalen 1971, p. 159). Whichever signal produces the largest cross-correlation with the received waveform is the one judged to have been delivered or, equivalently, the observer judges 's_0' if the difference in cross-correlations is below some criterion and 's_1' if above

$$\sum\sum\sum r_{xyt}s_{0xyt} - \sum\sum\sum r_{xyt}s_{1xyt}. \tag{4}$$

Note that the cross-correlation operation discussed here is quite different from the autocorrelation operation in the famous Reichardt (1961) model. One way of calculating cross-correlation is called matched filtering: the signal is passed through a linear filter that is matched to the stimulus (identical except reversed in time and space) and the output is sampled at some position and instant (Whalen 1971, pp. 168–170). The ideal observer model suggests that human motion performance is based upon neural spatio-temporal filters that are matched to the stimuli.

As Whalen (1971) showed, the performance of the ideal observer is given by

$$d' = \frac{\sqrt{\sum\sum\sum (s_{1xyt} - s_{0xyt})^2}}{\sigma}, \tag{5}$$

*Author for correspondence (wsi@gcal.ac.uk).

Received 12 September 2000 *Accepted* 27 November 2000

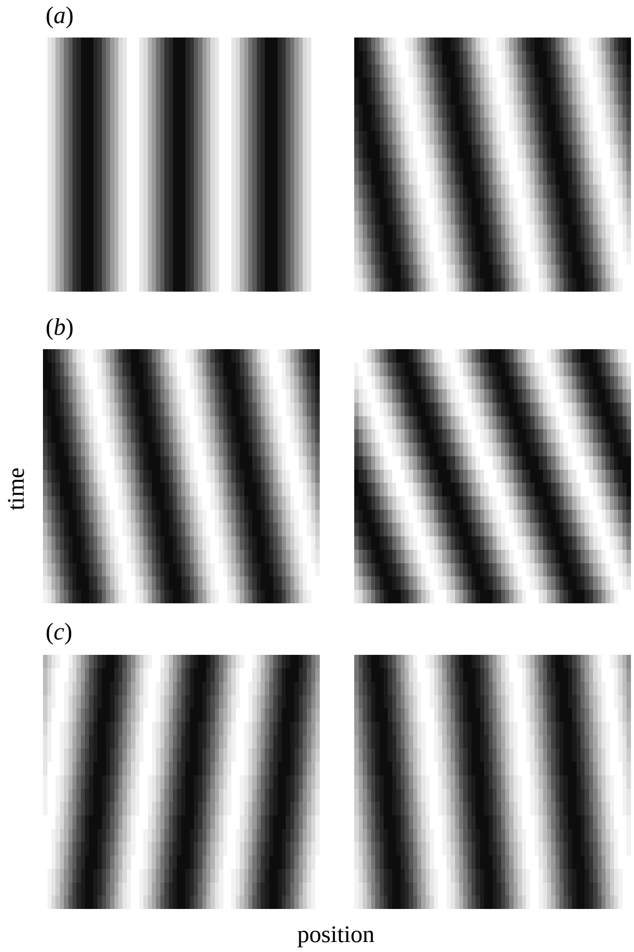

Figure 1. Space–time plots of the signal pairs to be discriminated in three tasks: (*a*) motion detection, (*b*) speed discrimination and (*c*) direction discrimination. The signals are vertical sine-wave gratings that are stationary or drift to the left or right.

where d' is the signal detection theory index of discriminability. Some insight into the model can be gained by rearranging equation (5). Define the average energy of the two signals as

$$E = \frac{1}{2}\sum\sum\sum(s^2_{0xyt} + s^2_{1xyt}), \tag{6}$$

and define the normalized correlation between the signals as

$$\rho = \frac{1}{E}\sum\sum\sum s_{0xyt}s_{1xyt}. \tag{7}$$

Expanding equation (5) and substituting equations (6) and (7) we obtain

$$d' = \frac{\sqrt{\sum\sum\sum(s^2_{0xyt} + s^2_{1xyt}) - 2\sum\sum\sum s_{0xyt}s_{1xyt}}}{\sigma}$$

$$= \frac{\sqrt{E - 2E\rho}}{\sigma} = \frac{\sqrt{2E(1-\rho)}}{\sigma}. \tag{8}$$

The discriminability of the two signals only depends upon their average energy E, their similarity or cross-correlation ρ and the standard deviation of the noise σ in which the signals are embedded. A real observer's performance might be described by a scaled version of equation (8), i.e.

$$d' = \frac{\sqrt{F2E(1-\rho)}}{\sigma} \tag{9}$$

where F represents the observer's efficiency (Burgess *et al.* 1981). One way in which real observers' efficiency might be lowered is if the filters they use are not precisely matched to the signals (in speed, spatial frequency, phase, size, location or duration). Other sources of inefficiency would include uncertainty about the stimulus parameters and additive internal noise (Burgess 1990).

We tested some predictions from the cross-correlator model in a series of three experiments. In the first experiment we presented drifting gratings having the same set of speed differences in three different motion tasks: motion detection, speed discrimination and direction discrimination. By varying the speed difference we vary ρ. When the data are plotted as a function of ρ, we expect all the data to fall on one curve. In the second experiment we varied ρ by changing the phases of the drifting gratings being discriminated. In the third experiment, we varied the contrast of the gratings.

3. METHODS

(a) *Methods common to all experiments*

The two observers, C.L. and J.M., had normal acuity.

The displays were presented on a Tektronix 608 oscilloscope with ^{15}P phosphor. The oscilloscope was controlled by a point-plotting memory buffer developed at the University of Alberta (Finley (1985) describes an older version). The system has 12 bits of control over pixel position and luminance. Viewing was binocular from a chin rest placed 82 cm from the face of the oscilloscope.

The stimuli were drifting sine-wave gratings having a spatial frequency of 3 cycles deg^{-1} and a mean luminance of 40 cd m^{-2}. The starting phase was fixed. The display on a given trial consisted of two drifting gratings, one above the other, with a fixation mark halfway between the two fields. The display was black (0.5 cd m^{-2}) outside the two stimulus fields. The two grating fields were 1.0° wide by 0.48° high. They were vertically separated by 0.04°. In the period between trials the grating fields were uniform patches with luminance of 40 cd m^{-2}.

The stimuli were constructed as movies consisting of 20 frames. The point plotter refreshed the display at a rate of 252 Hz. Each frame of the movie was displayed for three screen refreshes, so the frame duration was 11.89 ms. The stimuli were drifting gratings with added dynamic normal noise, plotted in a regular raster pattern (each field was composed of 66 × 31 dots with an interdot separation of 0.015°). The noise was generated by a normal pseudo-random number generator (using Marsaglia's polar method on uniform variates from a multiply-with-carry generator) and added independently onto each plotted point.

Two stimulus fields containing moving gratings were presented in each trial. The top field contained the standard stimulus and the bottom field contained either the standard or the comparison with equal probability. The subject's task was to judge whether the bottom field contained the standard or the comparison. Responses were made by a button box and a beep sounded if an incorrect response was made. One standard and one comparison were discriminated within a block of 100 trials.

Between three and five blocks of trials were run for each stimulus pairing.

(b) *Speed difference experiment*

There were three conditions: motion detection, speed discrimination and direction discrimination. In each case the observer discriminated between a standard and a comparison moving sine-wave grating in noise. The standard and comparison were alike in all ways (contrast, spatial frequency and phase) except speed. In motion detection the standard stimulus was stationary and the comparison speeds were 0.17, 0.34, 0.51, 0.67 and $0.84^\circ\ s^{-1}$. In speed discrimination the standard stimulus drifted at $0.84^\circ\ s^{-1}$ and the comparison speeds were 1.01, 1.18, 1.35, 1.51 and $1.68^\circ\ s^{-1}$. In direction discrimination the standard and comparison stimuli drifted left and right, respectively, at speeds of 0.08, 0.17, 0.25, 0.34 and $0.42^\circ\ s^{-1}$. Note that all conditions used the same set of speed differences between the standard and comparison. The various conditions were presented in random order.

The signal contrast was 75% and the noise contrast standard deviation was 12.5% (noise clipped at two standard deviations).

(c) *Phase difference experiment*

The stimuli were drifting gratings in noise. The task was direction discrimination. The speed difference was fixed at $0.28^\circ\ s^{-1}$, and the two signals differed in phase by 0, $\pi/4$, $\pi/2$, $3\pi/4$ and π radians. The starting phase was fixed. The signal contrast was 5% and the noise standard deviation contrast was 30%.

(d) *Contrast response experiment*

The stimuli were gratings embedded in dynamic noise that moved to the left or right. Their speed was $\pm 0.28^\circ\ s^{-1}$ and the task was direction discrimination. The signal contrast had values of 5, 12.5, 25, 50 and 75% and the noise standard deviation contrast was 12.5% (noise clipped at two standard deviations).

4. RESULTS

(a) *Speed difference experiment*

In this experiment we measured the observers' ability to discriminate between two signals s_0 and s_1 as a function of their speed difference. Three conditions were run: motion detection, speed discrimination and direction discrimination. To an ideal observer all that matters is the difference in the speeds of the two signals; the speed difference affects the cross-correlation of the signals. The smaller the speed difference, the more similar the signals and the poorer the observer's ability to discriminate them. In terms of the space–time diagrams in figure 1, the smaller the speed difference the more similar the orientations of the signals in space–time. If these orientations are identical the cross-correlation is unity, whereas if the orientations are at right angles the cross-correlation is zero.

The data are plotted in figure 2. The curves shown are least-squares fits of equation (9) to all the data grouped together. The fitted curve shows the performance of an inefficient cross-correlator. The 95% confidence intervals for the observers' efficiencies were $0.000\,20 \pm 0.000\,04$ (C.L.) and $0.000\,27 \pm 0.000\,04$ (J.M.). Separate fits to the motion detection, speed discrimination and direction discrimination data produced parameter estimates that were not significantly different from the grouped fit

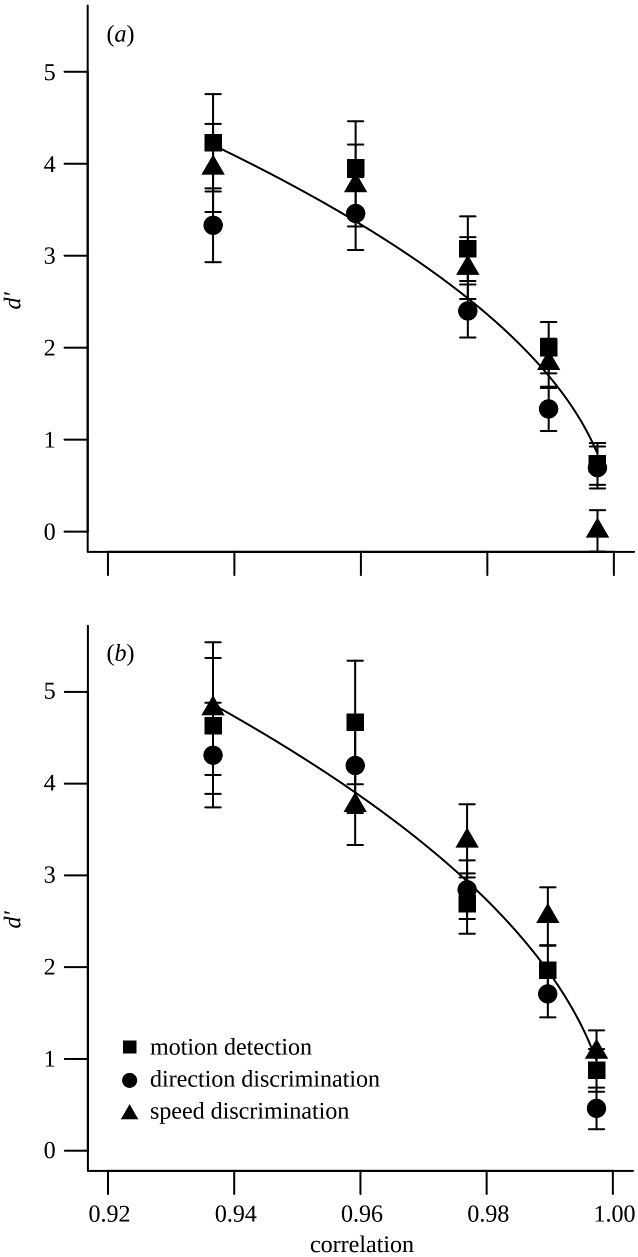

Figure 2. d' as a function of the correlation between the signals being discriminated in the motion detection, speed discrimination and direction discrimination tasks for observers (*a*) C.L. and (*b*) J.M. The signals differed in speed. The least-squares fits of equation (9) to all the data are shown. The error bars show 95% confidence intervals.

(Gallant 1987, p. 59). As can be seen, equation (9) fits the data quite well; the r^2 (defined as $1 - (SS_{error}/SS_{total})$) for J.M. is 0.93 and that for C.L. 0.89.

We predicted that the data would be described by a cross-correlator model. This prediction was borne out, although our observers were quite inefficient. Studies with static grating detection have found an efficiency of about 0.5 (Burgess *et al.* 1981; Burgess & Ghandeharian 1984). We think the discrepancy was due to the fact that we were dealing with discrimination of suprathreshold patterns rather than detection of near-threshold patterns. The ideal observer's performance in a grating detection task is governed only by the signal energy and the noise level. Inefficiency can be associated with each of these:

the signal energy can be extracted poorly (sampling efficiency) and internal noise can be added to the noise present in the stimulus. In a discrimination task a third factor enters into the picture: the cross-correlation between the signals. Inefficiency can be associated with the cross-correlation in the following way. Cross-correlation can be implemented as matched filtering and sampling inefficiency will result if the observer uses filters that are not precisely matched to the signals. Furthermore, those filters used by the observer may be more correlated than the signals. Thus, the observer may act as though the signals are more similar than they really are, giving low efficiency. This explains how it is possible for highly detectable signals to be poorly discriminated.

A second prediction from the cross-correlator model was that only the correlation between the two signals being discriminated mattered and that one curve would fit the results for all motion tasks. This prediction was also supported by the data.

Although the results were consistent with the cross-correlator model, they were also consistent with the Adelson & Bergen (1985) model. Fig. 9 of Adelson & Bergen depicts the following model:

$$[r(x,t) * h_0(x,t)]^2 + [r(x,t) * h_1(x,t)]^2, \tag{10}$$

where h_0 and h_1 are orientated spatio-temporal filters in quadrature phase and the asterisks represent convolution. The term 'motion energy' is a misnomer because the output of this model is a function of space and time whereas energy is a single number. (The energy in the bands passed by the filters could be obtained by integrating the output over space and time.) Suppose that the observers base their decision on an instantaneous sample of the output. The resulting scheme is known as a quadrature receiver in the signal detection literature (Whalen 1971, pp. 196–209). A quadrature receiver is the optimal way of detecting or discriminating waveforms the phase of which is unknown (or random).

The Adelson & Bergen model was modified in order to obtain quantitative predictions for our experiments. The spatio-temporal filters were specified as sine- and cosine-phase filters matched to the signals, giving the model

$$\begin{aligned}&\left\{\left[\sum\sum\sum r_{xyt}\cos(2\pi fx + v_0 t)\right]^2\right.\\ &\left.+\left[\sum\sum\sum r_{xyt}\sin(2\pi fx + v_0 t)\right]^2\right\}\\ &-\left\{\left[\sum\sum\sum r_{xyt}\cos(2\pi fx + v_1 t)\right]^2\right.\\ &\left.+\left[\sum\sum\sum r_{xyt}\sin(2\pi fx + v_1 t)\right]^2\right\}.\end{aligned} \tag{11}$$

Note that we have used multiplication by cosine and sine waveforms combined with summation in place of filtering. The scheme represented by equation (11) is equivalent to the quadrature receiver (Whalen 1971, pp. 196–209). The quadrature receiver cross-correlates the received waveform with sine- and cosine-phase versions of the expected signals and sums the squared results. Whichever signal produces the best match with the received waveform (as computed by the quadrature receiver mechanism) is judged to be the one delivered.

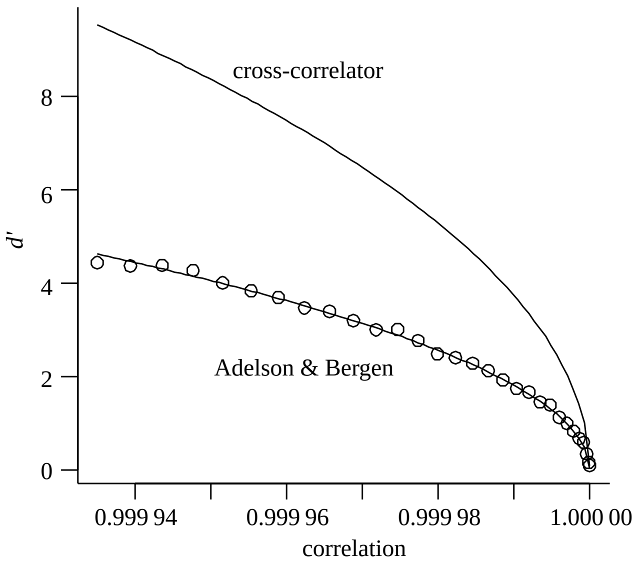

Figure 3. Predicted performance (d') for an ideal cross-correlator and an Adelson & Bergen (1985) sensor as a function of the cross-correlation between the signals being discriminated. The Adelson & Bergen data points were obtained through simulation; the curve is a least-squares fit of equation (9).

The performance of the quadrature receiver has been analysed for one-dimensional signals (Whalen 1971, pp. 202–205). The results are somewhat complicated and we computed the model's discrimination performance through simulation. The cross-correlator and the quadrature receiver are compared in figure 3. The simulated performance of the Adelson & Bergen sensor is well-fitted by equation (9); thus it behaves like an inefficient cross-correlator. Since both humans and the Adelson & Bergen sensor act like inefficient cross-correlators, we cannot tell from this experiment whether humans use a cross-correlator or quadrature receiver mechanism. In the phase difference experiment we tested which of the two mechanisms underlies human performance.

(b) ***Phase difference experiment***

The results of the speed difference experiment were consistent with the cross-correlator and the Adelson & Bergen models. We now tested between the two mechanisms in the following way. The Adelson & Bergen sensor is inefficient due to its loss of phase information, which is a consequence of the squaring in equation (11). The loss of phase means that the Adelson & Bergen sensor cannot discriminate between two moving gratings that are identical except in phase. However, the cross-correlator can use the phase; the greater the phase difference between the two signals, the smaller the cross-correlation and the greater the discriminability.

The data are plotted in figure 4. The solid curves shown are the least-squares fits of equation (9). The 95% confidence intervals for the efficiency were 0.0057 ± 0.0013 for J.M. and $0.0060 \pm 0.000\,90$ for C.L. The fit of the inefficient cross-correlator model was satisfactory; the r^2 for J.M. was 0.94 and that for C.L. 0.98. The dashed lines show the performance of the Adelson & Bergen sensor obtained by 5000 simulation trials.

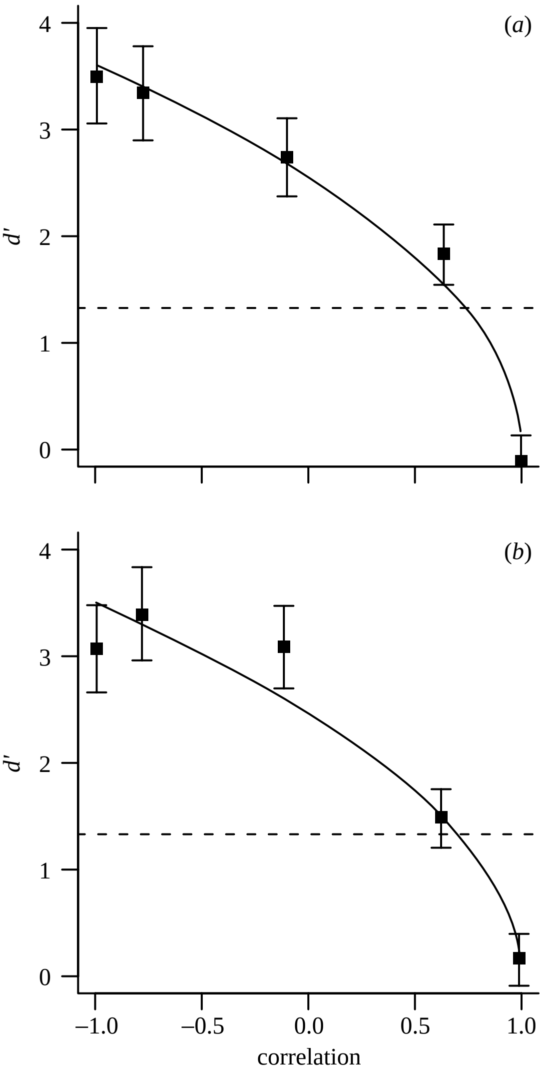

Figure 4. d' as a function of the correlation between the signals being discriminated in a direction discrimination task for observers (*a*) C.L. and (*b*) J.M. The error bars show 95% confidence intervals. The phase difference between the signals was varied and the speed difference was fixed. The least-squares fits of equation (9) are shown. The dashed lines show the performance of the Adelson & Bergen model.

As figure 4 shows, the discriminability of the two moving gratings is enhanced as their phases become more different. This pattern is consistent with the cross-correlator model and the quantitative fit of the model to the data is good. The Adelson & Bergen sensor's output is not affected at all by the phase of the input; therefore the data are not consistent with an Adelson & Bergen motion-sensing mechanism.

We have demonstrated that human observers, like the ideal cross-correlator, can use phase information for discriminating moving gratings. The same result was obtained by Burgess & Ghandeharian (1984) for static gratings. This is not to say that humans never use a phase-insensitive mechanism in motion discrimination. Indeed, if the phase of the gratings in the experiment were randomized we would expect observers to use a version of the quadrature receiver (a modified Adelson & Bergen sensor), which is the ideal observer when phase is unknown. Seemingly trivial changes in the experimental set-up can completely change the ideal observer and we would expect human performance to mirror these changes.

It is traditional in the motion literature (McKee & Watamaniuk 1994) to consider motion and position information to be extracted by different mechanisms. According to this way of thinking, the phase difference experiment taps a position mechanism rather than a motion mechanism. This idea has some problems with it. First, what is the detailed formulation of the position mechanism? We suspect it would be cross-correlation under a different name. Second, how are the two mechanisms' outputs combined? In the cross-correlator approach, motion and phase are combined in a completely automatic and ideal way.

(c) *Contrast response experiment*

Although the data from the speed difference and the phase difference experiments were consistent with the cross-correlator model, the model makes a prediction about the dependence of motion sensing on stimulus contrast that is at odds with previous research. Equation (8) tells us that the function relating d' and stimulus contrast is a line through the origin (since the square-root of the energy is proportional to the contrast). However, Nakayama & Silverman (1985) found that, for sine-wave gratings, the detection of a sudden jump did not improve any further once the contrast exceeded *ca.* 2%. We measured direction discrimination as a function of contrast in order to test whether the contrast response was approximately linear (as predicted) or sharply saturating.

Figure 5 shows d' as a function of grating contrast for both observers. The predicted function for a cross-correlator is a line through the origin. Although a line through the origin does a reasonable job of describing the data, the true function appears to saturate. In order to test for the presence of saturation we fitted a power curve to the data by least squares. The 95% confidence intervals for the fitted exponent were 0.68 ± 0.08 for J.M. and 0.53 ± 0.05 for C.L. Neither of these confidence intervals includes an exponent value of 1.0, which gives the straight line predicted for the cross-correlator. Therefore, we conclude that the direction discrimination mechanism shows mild saturation.

We have been arguing that motion detection, speed discrimination and direction discrimination can be explained by a cross-correlator model. The cross-correlator can be implemented using spatio-temporal matched filtering. Directionally selective simple cells in the striate cortex perform the necessary spatio-temporal filtering (Hamilton *et al.* 1989; McLean & Palmer 1989, 1994; McLean *et al.* 1994; Reid *et al.* 1997). Thus, we are proposing that the human performance in our motion tasks was due to directionally selective simple cells. It turns out that the contrast response of V1 simple cells does saturate, but not sharply (Sclar *et al.* 1990). The semi-saturation constant varies widely between cells—the distribution is almost uniform between the contrasts of 0 and 100%, with a median value of 33%. Thus, our psychophysical finding of gentle saturation is consistent with the performance of V1 cells. Further psychophysical support comes from the fact that perceived speed varies linearly with contrast over the full range of contrasts

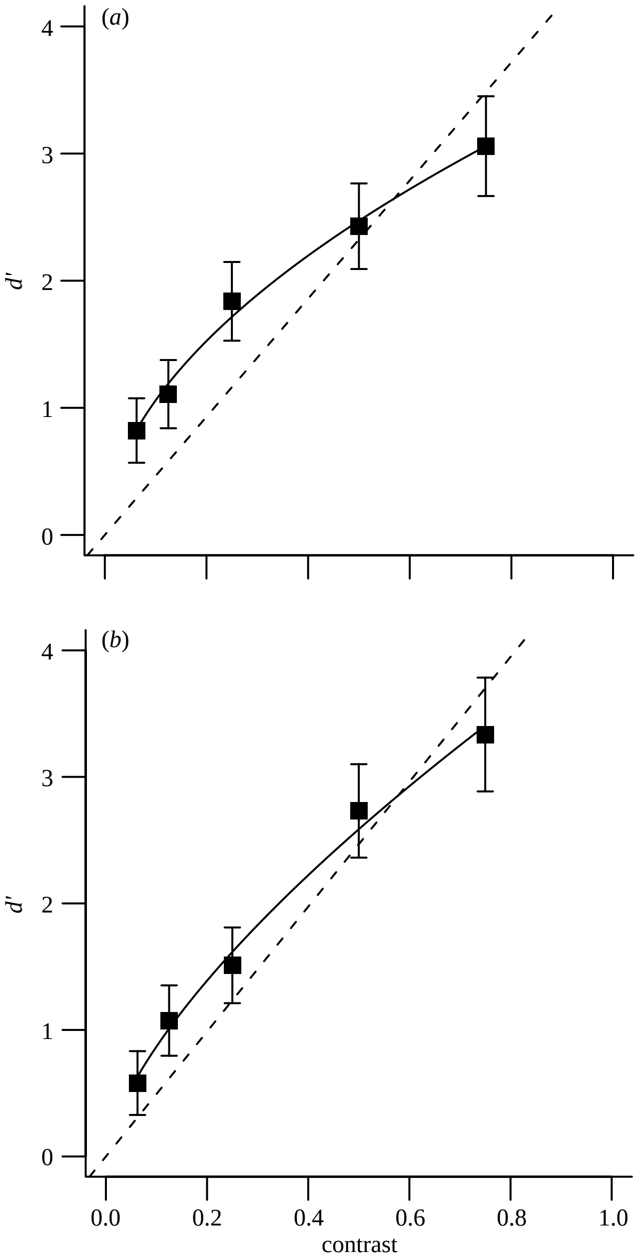

Figure 5. d' (with 95% confidence intervals) for discriminating the direction of two gratings as a function of their contrast for observers (*a*) C.L. and (*b*) J.M. Each dashed line through the origin shows the cross-correlator prediction; each solid curve is a power-law fit with exponent less than unity.

(Thompson 1982; Stone & Thompson 1992; Hawken *et al.* 1994; Gegenfurtner & Hawken 1996).

Why do our data not agree with those of Nakayama & Silverman (1985)? The stimuli, methods and tasks were very similar in the two experiments. One difference is that our display contained luminance noise, whereas Nakayama & Silverman's display did not. Spekreijse & Van der Tweel (1965) found that adding luminance noise linearizes the visual evoked response and perhaps adding noise linearized the contrast response in our case. Another difference is that Nakayama & Silverman's motion display used two frames, whereas ours used 20. It is possible that our multiframe stimulus was better tailored to V1 simple cells. Perhaps the response of MT cells to the two-frame display was greater and, thus, MT cells governed the observers' data. MT cells tend to show steep contrast response functions; the distribution has an exponential shape with most cells having very small semi-saturation constants with a median value of 7% (Sclar *et al.* 1990). Thus, the performance of MT cells is consistent with the psychophysical data gathered by Nakayama & Silverman.

5. DISCUSSION

We performed three experiments aimed at testing whether human motion perception was similar to that of an ideal mechanism that cross-correlates the noisy stimulus received with each of the expected spatio-temporal signals. In the first experiment we varied the speed difference of the moving gratings to be discriminated and thereby varied their cross-correlation. The ability of the observers to discriminate the gratings declined as the cross-correlation of the standard and comparison increased and the functional form of this relationship was as predicted by the cross-correlator model (figure 2). The data in the first experiment were also consistent with the Adelson & Bergen (1985) motion sensor, since this sensor also behaves like an inefficient cross-correlator (figure 3). In the second experiment we tested between the alternative models by measuring direction discrimination as a function of the initial phase difference. Varying the phase difference changes the cross-correlation between the standard and comparison and so an ideal cross-correlator will perform in a similar way to that seen when the speed difference is manipulated. However, the Adelson & Bergen sensor is insensitive to phase and so performance will be unaffected by the phase difference. The human observers' performance was affected by the phase difference in the way predicted from the cross-correlator model (figure 4). In the third experiment we checked another prediction from the cross-correlator model: the cross-correlator's performance increases linearly as contrast increases. We thought it was necessary to check this prediction since Nakayama & Silverman (1985) found that performance saturated once contrast exceeded *ca.* 2%. The performance of our observers increased with contrast in an almost linear way (though gently saturated) (see figure 5).

To summarize, the psychophysical data from the experiments are consistent with the cross-correlator model. The next question concerns the physiological substrate of the psychophysics. A cross-correlator mechanism can be implemented as a matched filter. That is, the observer can cross-correlate the noisy waveform received with each of the expected signals by passing the waveform through filters matched to these signals. In our case the waveforms being cross-correlated were oriented gratings in space–time. It so happens that the visual system contains filters that will serve nicely as matched filters: V1 directionally selective simple cells. These cells have the correct spatio-temporal receptive fields for doing the job (Hamilton *et al.* 1989; McLean & Palmer 1989, 1994; McLean *et al.* 1994; Reid *et al.* 1997) and, to a large degree, they act as linear filters (Hamilton *et al.* 1989; Reid *et al.* 1991, 1997; McLean & Palmer 1989, 1994; Jagadeesh *et al.* 1993; McLean *et al.* 1994), and their contrast response curves do not saturate sharply (Sclar *et al.* 1990).

We propose that V1 directionally selective simple cells are the physiological substrate for motion detection and discrimination for gratings. The observer monitors the outputs of two such cells, one tuned to the standard and one tuned to the comparison. The response to each stimulus depends on which cell produces the bigger output (standard or comparison). This idea that a small population of cells suffices for explaining motion discrimination was also put forward by Newsome *et al.* (1989) in their work on MT cells.

Our suggestion that human motion detection and discrimination use a matched filter mechanism is compatible with known electrophysiology. It is also compatible with previous psychophysical results. Burr *et al.* (1986) used a masking approach in order to reveal the spatio-temporal filters underlying motion detection. The filters shown in their figure 6 are precisely the sort that we believe were used by our observers. Watson & Turano (1995) measured the contrast energy threshold for discriminating the direction of a drifting grating (with Gaussian envelope). Their figure 11 shows a space–time image of the most readily detected moving stimulus and Watson & Turano interpreted this image as showing the most sensitive spatio-temporal receptive field underlying motion detection. These authors pointed out that, if the visual system uses matched filtering, the optimal motion stimulus corresponds to the matched filter. Reisbeck & Gegenfurtner (1999) measured spatio-temporal discrimination contours that were orientated in spatial and temporal frequency (see their fig. 3), indicating the presence of oriented spatio-temporal filters of the type we require.

In conclusion, we have reported evidence that supports the idea that human motion detection and discrimination use simple linear matched filtering. This matched filtering can be achieved by passing the noisy received stimulus through spatio-temporal filters that are oriented in space–time. Other authors have proposed such a first stage in more complicated motion models (Adelson & Bergen 1985; Watson & Ahumada 1985). However, our data are well explained by simple matched filtering; indeed we found that human motion sensing is phase sensitive and so the extra processing proposed by Adelson & Bergen (1985) is not carried out when phase information is available. We propose that human motion detection and discrimination of gratings is based on the outputs of V1 directionally selective simple cells, which act as linear spatio-temporal filters orientated in space–time.

We thank Claire Louden and Jane McGhie for their help in collecting the data.

REFERENCES

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

Burgess, A. E. 1990 High level visual decision efficiencies. In *Vision: coding and efficiency* (ed. C. Blakemore), pp. 431–440. New York: Cambridge University Press.

Burgess, A. E. & Ghandeharian, H. 1984 Visual signal detection. I. Ability to use phase information. *J. Opt. Soc. Am.* **1**, 900–905.

Burgess, A. E., Wagner, R. F., Jennings, R. J. & Barlow, H. B. 1981 Efficiency of human visual signal discrimination. *Science* **214**, 93–94.

Burr, D. C., Ross, J. & Morrone, M. C. 1986 Seeing objects in motion. *Proc. R. Soc. Lond.* B **227**, 249–265.

Finley, G. 1985 A high-speed plotter for vision research. *Vision Res.* **25**, 1993–1997.

Gallant, A. R. 1987 *Nonlinear statistical models*. New York: Wiley.

Gegenfurtner, K. R. & Hawken, M. J. 1996 Perceived speed of luminance, chromatic, and non-Fourier stimuli: influence of contrast and temporal frequency. *Vision Res.* **36**, 1281–1290.

Hamilton, D. B., Albrecht, D. G. & Geisler, W. S. 1989 Visual cortical receptive fields in monkey and cat: spatial and temporal phase transfer function. *Vision Res.* **29**, 1285–1308.

Hawken, M. J., Gegenfurtner, K. R. & Tang, C. 1994 Contrast dependence of colour and luminance motion mechanisms in human vision. *Nature* **367**, 268–270.

Jagadeesh, B., Wheat, H. S. & Ferster, D. 1993 Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. *Science* **262**, 1901–1904.

McKee, S. P. & Watamaniuk, S. N. J. 1994 The psychophysics of motion perception. In *Visual detection of motion* (ed. A. T. Smith & R. J. Snowden), pp. 85–114. New York: Academic Press.

McLean, J. & Palmer, L. A. 1989 Contribution of linear spatio-temporal receptive field structure to velocity selectivity of simple cells in area 17 of cat. *Vision Res.* **29**, 675–679.

McLean, J. & Palmer, L. A. 1994 Organization of simple cell responses in the three-dimensional (3-D) frequency domain. *Visual Neurosci.* **11**, 295–306.

McLean, J., Raab, S. & Palmer, L. A. 1994 Contribution of linear mechanisms to the specification of local motion by simple cells in areas 17 and 18 of the cat. *Visual Neurosci.* **11**, 271–294.

Nakayama, K. & Silverman, G. H. 1985 Detection and discrimination of sinusoidal grating displacements. *J. Opt. Soc. Am.* A **2**, 267–274.

Newsome, W. T., Britten, K. H. & Movshon, J. A. 1989 Neuronal correlates of a perceptual decision. *Nature* **341**, 52–54.

Reichardt, W. 1961 Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In *Sensory communication* (ed. W. A. Rosenblith), pp. 303–317. Cambridge, MA: MIT Press.

Reid, R. C., Soodak, R. E. & Shapley, R. M. 1991 Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. *J. Neurophysiol.* **66**, 505–529.

Reid, R. C., Victor, J. D. & Shapley, R. M. 1997 The use of m-sequences in the analysis of visual neurons: linear receptive field properties. *Visual Neurosci.* **14**, 1015–1027.

Reisbeck, T. E. & Gegenfurtner, K. R. 1999 Velocity tuned mechanisms in human motion processing. *Vision Res.* **39**, 3267–3285.

Sclar, G., Maunsell, J. R. & Lennie, P. 1990 Coding of image contrast in central visual pathways of the macaque monkey. *Vision Res.* **30**, 1–10.

Spekreijse, H. & Van der Tweel, L. H. 1965 Linearization of evoked responses to sine wave-modulated light by noise. *Nature* **205**, 913–915.

Stone, L. S. & Thompson, P. 1992 Human speed perception is contrast dependent. *Vision Res.* **32**, 1535–1549.

Thompson, P. 1982 Perceived rate of movement depends on contrast. *Vision Res.* **22**, 377–380.

Watson, A. B. & Ahumada Jr, A. J. 1985 Model of human visual-motion sensing. *J. Opt. Soc. Am.* A **2**, 322–341.

Watson, A. B. & Turano, K. 1995 The optimal motion stimulus. *Vision Res.* **35**, 325–336.

Whalen, A. D. 1971 *Detection of signals in noise*. New York: Academic Press.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.