Pergamon

0042-6989(94)00258-4

Vision Res. Vol. 35, No. 12, pp. 1771–1783, 1995

0042-6989/95 $9.50 + 0.00

Perceived Motion of Contrast-modulated Gratings: Predictions of the Multi-channel Gradient Model and the Role of Full-wave Rectification

A. JOHNSTON,* C. W. G. CLIFFORD*

Received 28 January 1994; in revised form 11 May 1994; in final form 17 September 1994

The paper examines the perception of motion in contrast-modulated sine-wave grating patterns. These non-rigid motion patterns give rise to a spatially-structured motion percept in which perceived speed varies with spatial position. We measured the perceived motion of the low contrast regions of amplitude-modulated gratings as a function of the carrier frequency, the carrier speed, the shape of the modulation signal and the modulation depth. We found that for static carriers perceived speed was greatest in the low contrast regions of the display. The speed of the low contrast regions was underestimated and perceived speed decreased as the spatial frequency of the carrier increased. When the direction of the motion of the carrier was opposite to that of the contrast modulation, the low contrast regions could appear to be stationary. The perceived speed of the contrast modulation increased with modulation depth. The brightness contrast of the carrier grating had little effect on perceived speed of contrast-modulated patterns for average contrasts of over 10%. A motion model which had full-wave rectification as an explicit pre-processing stage followed by low-pass filtering or some other selection criterion, would predict that the motion of contrast-modulated gratings should appear rigid and that the motion of the envelope should be judged correctly. The Multi-channel Gradient Model however predicts both the structured motion field experienced when viewing these second-order motion patterns and the reductions in perceived speed as a function of carrier spatial frequency and carrier speed.

Motion Second-order motion Contrast modulation Rectification Computational model

INTRODUCTION

There is general agreement that the analysis of spatial pattern and image motion by the human visual system involves an initial linear stage in which the image is processed by a bank of spatio-temporal filters. This consensus leaves questions about how to characterize the filters, how the outputs of the linear stage can be combined to recover properties of the image and what the functional role of the filters might be as components of computational schemes for pattern and motion analysis.

The motion of one-dimensional spatial patterns can be represented in an image showing the changes in brightness over space and time [Fig. 1(a)]. In a space–time image a spatial pattern moving rigidly at a constant speed generates oriented contours along which brightness is constant. Image speed is given by the orientation of the contours. The Fourier transform of the space–time image is also oriented [Fig. 1(b)] and is centred on the origin reflecting the fact that the ratio of the temporal frequency to spatial frequency of any component is constant for rigid motion, the constant factor indicating the image speed. The Motion From Fourier Components (MFFC) approach attempts to measure image speed by recovering the orientation through the origin of energy in the frequency domain (Adelson & Bergen, 1985; Chubb & Sperling, 1988; Heeger, 1987).

However, some space–time patterns which give rise to motion percepts do not have Fourier transforms characteristic of the motion of rigid patterns. Here we consider one class of second-order patterns, contrast-modulated sine gratings (Derrington & Badcock, 1985; Turano & Pantle, 1989), in which the contrast of a static sine-wave luminance grating is dictated by the value of a drifting sine-wave modulation signal [Fig. 1(d)]. The Fourier transform of this stimulus [Fig. 1(e)] shows that the spatio-temporal energy has a local orientation which is shifted away from the origin to a degree which is dependent upon the spatial frequency of the carrier

*Department of Psychology, University College London, Gower Street, London WC1E 6BT, England [*Email* alan@psychol.ucl.ac.uk].

grating. Fleet and Langley (1994) provide a detailed analysis of this class of non-rigid motion.

Since it is difficult to account for the motion of these patterns and other "non-Fourier" motion stimuli on a MFFC basis, it has been suggested that we have multiple motion systems; a MFFC system and an additional system which applies a motion energy analysis after band-pass filtering and full-wave rectification (Chubb & Sperling, 1988, 1989; Kim & Wilson, 1993; Wilson, Ferrera & Yo, 1992). We can include a stage of low-pass filtering (Sperling, 1989), band-pass filtering 1 octave lower than the carrier (Wilson *et al.*, 1992) or band-pass filtering at the modulation frequency (Zhou & Baker, 1993) in the non-Fourier route to ensure that the replicates introduced in the Fourier transform of the signal by the squaring or rectification operation are eliminated (see also Derrington, 1987, 1990; Derrington, Badcock & Henning, 1993; Turano & Pantle, 1989). This process would ensure a first-order motion signal which can then be processed in the usual way. The outputs from the Fourier and non-Fourier pathways would also need to be combined (Sperling, 1989). Figure 2 summarizes this view.

At the heart of the Fourier vs non-Fourier distinction and the categorization of first-order vs second-order is the consideration of whether the Fourier transform of the motion pattern lies on a plane through the origin or not. This is really a condition of rigid vs non-rigid motion. Thus the perceived role of the non-Fourier route is to recover from what is a non-rigid motion signal, an influence which is everywhere present and uniform which we call the second-order or non-Fourier motion signal. Rectification followed by low-pass filtering, or some other selection criteria, should allow the speed of the contrast modulation to be seen veridically at all points in the image [Fig. 1(g, h)]. But do non-rigid motion patterns appear to move rigidly and is the speed of a contrast modulation reported correctly?

In an earlier paper (Johnston, McOwan & Buxton, 1992) we indicated that the motion field of a

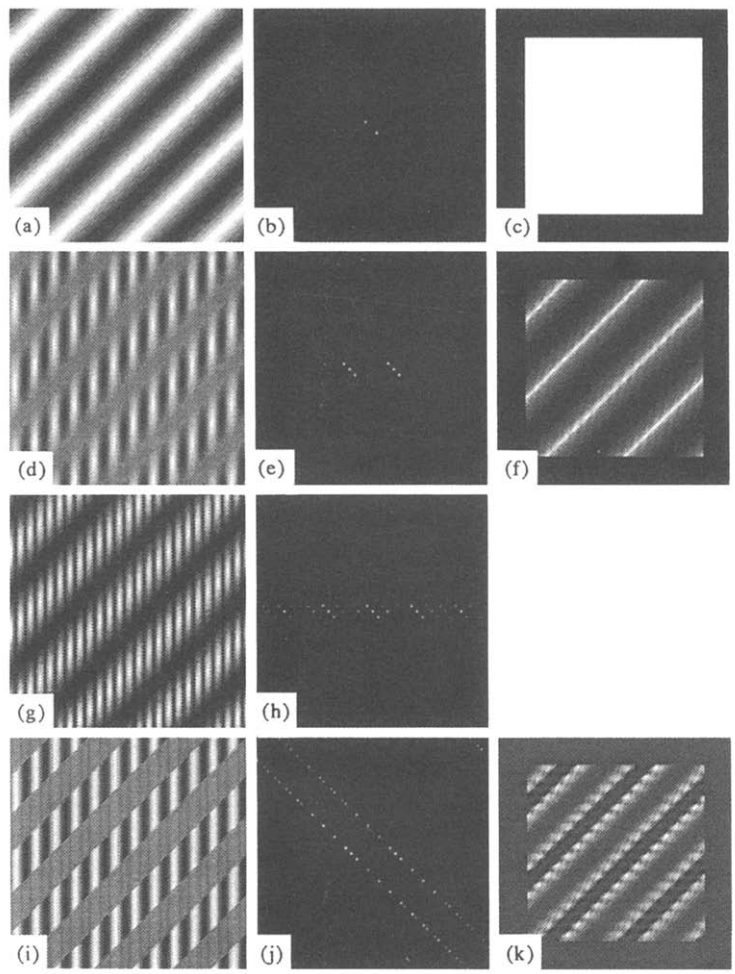

FIGURE 1. (a) A space–time image showing a drifting grating. Time increases down the figure. (b) The Fourier transform of (a) is also oriented. (c) The results of the McGM v.3. By convention the border is set to zero and then the output is scaled to the full range. The output is unstructured, motion is computed correctly at all points in the image. (d) A space–time image showing a contrast-modulated sine grating. (e) The Fourier transform of (d) showing motion energy the orientation of which depends upon the modulation speed and the position of which depends upon the spatial frequency of the carrier grating. (f) The results of the McMG v.3. Motion to the left is brighter than the boundary and motion to the right is darker than the boundary. Note that the output is spatially structured, therefore the model does not signal the motion of the contrast signal *per se*. (g) A rectified contrast-modulated grating. (h) The Fourier transform has energy through the origin. (i) Square-wave contrast modulation and its transform (j). (k) The model signals motion at the discontinuities.

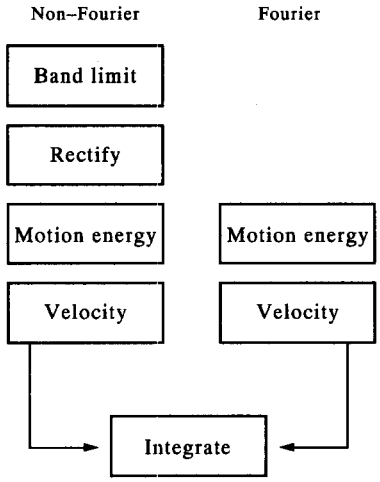

FIGURE 2. The two-mechanism model. The motion of first-order signals is seen by the Fourier pathway and second-order motion signals are seen by the non-Fourier pathway.

contrast-modulated grating appeared spatially structured. Observers saw the motion of the modulation in the low-contrast regions of the pattern and no motion in the high contrast regions. An interesting parallel is that in the space–time image the spatial orientation of the modulation is only seen in the low contrast regions. In the high contrast regions the image appears to be oriented vertically. This basic structure was mirrored in the output of the Multi-channel Gradient Model (McGM) (Johnston & Clifford, 1995; Johnston *et al.*, 1992).

Consider the motion of one-dimensional spatial patterns. Gradient models compute motion [equation (1)] as the ratio of the partial derivatives of image brightness with respect to space and time (Fennema & Thompson, 1979; Heeger & Simoncelli, 1995; Horn & Schunck, 1981; Johnston *et al.*, 1992; Sobey & Srinivasan, 1991; Verri, Straforini & Torre, 1992; Young & Lesperance, 1993)

$$v = -\frac{\partial I/\partial t}{\partial I/\partial x}. \quad (1)$$

However, gradient schemes are thought to be unreliable because the quotient becomes infinite when the denominator takes values near zero resulting in errors close to points in the image where we find peaks, troughs and points of inflection with zero gradients in image brightness. The McGM is motivated by the problem of conditioning the quotient. When the first derivative of image brightness is zero it is unlikely that the second- and higher-order derivatives are also zero. If we assume that speed is constant within our region of interest, we can get an estimate of image speed by forming a ratio of the partial derivatives of any derivative measure with respect to space and time. These measures are then combined using a standard least-squares formulation. In the least-squares quotient the denominator is a sum of squared terms and therefore can only be zero when all of the derivatives on the denominator are zero. The filter outputs are combined by non-linear product and squaring operations and any physical implementation of the algorithm would include components whose response to sine gratings would resemble a full-wave, squared rectification of the input signal. However, it is important to note that the function of rectification in the McGM model is to condition the calculation of a quotient not to provide an alternative pre-processing stage prior to motion energy analysis. It should be clear that these approaches are not mathematically equivalent. A full description of the current version of the model (McGM v.3) is presented in the Appendix of Johnston and Clifford (1995).

We reported in an earlier paper (Johnston *et al.*, 1992) that the output of the McGM is qualitatively correct for contrast-modulated patterns with static carriers in that it provides a spatially-structured field in which speed is maximal in the low contrast regions. In this paper we examine the quantitative predictions of the model by measuring psychophysically the perceived speed of the low contrast regions of contrast-modulated gratings and comparing the psychophysical data with the predictions of the model.

METHOD

Apparatus and stimuli

Space–time images were constructed using a SUN Sparcstation II workstation and then passed to a PC-based graphics system which acted as a slave to the workstation. The image sequences were displayed on a carefully gamma-corrected Manitron monitor driven by a Matrox graphics board. The graphics board delivered 8 bits per pixel to give 256 levels of grey. Animation was achieved by drawing a luminance ramp on the display and then passing the data from the space–time image to the colour look-up table frame by frame. Gamma correction and manipulation of image speed was accomplished at the generation stage. The frame rate was 59.5 Hz and the mean luminance of each stimulus was 17 cd/m^2. The display was viewed in a darkened room. Contrast was matched between luminance and contrast domain stimuli by ensuring that they had the same mean deviation of luminance from the mean. The contrast-modulated stimuli were generated using the equation

$$I(x,t) = \mu(1 + A(1 + M \times \cos(k_m x - \omega_m t + \phi)) \cos(k_c x - \omega_c t)) \quad (2)$$

where μ is the mean luminance, A is the contrast of the carrier, M is the modulation depth ($0 \leqslant M \leqslant 1$), k_m and ω_m are the spatial and temporal frequencies of the modulation, k_c and ω_c are the spatial and temporal frequencies of the carrier grating and ϕ is the initial phase of the modulation. Thus, for a modulation depth of 1.0, the peak to peak variation in intensity for contrast-modulated gratings was twice that for the corresponding luminance gratings. The contrast of the sine grating was always 0.25 unless otherwise stated.

The stimuli were presented in a pair of circular windows, each subtending 2.09 deg of visual angle, on a uniform 17 cd/m^2 background. The subject was required to fixate a prominent dark fixation spot placed midway

between the centres of the two windows. The visual angle between the direction of fixation and the centre of each window was 1.42 deg. The viewing distance was 2.0 m.

Each stimulus presentation was preceded by a bright warning frame which served to maintain a uniform level of light adaptation. The stimuli were multiplied by a linear temporal ramp at onset and offset. Onset time, initial phase and stimulus duration were randomized to prevent the use of spatial translation as a basis for decisions. Onset time varied over a 170 msec range. Stimulus duration varied between 1.82 and 1.98 sec, with the ramps comprising the first and last 80 msec.

Procedure

Observers compared the perceived speed of a sine grating and the low contrast or high contrast regions of the envelope of a contrast-modulated grating using an adaptive method of constants paradigm (Watt & Andrews, 1981), in which the speed of the sine grating was varied systematically from trial to trial to determine a psychometric function. Subjects typically compared the speed of the low contrast regions with the speed of the bright regions of the grating. The test and standard stimuli were randomly allocated to the left and right windows on each trial, and an interval of at least 5 sec was maintained between trials, minimizing the effects of motion adaptation. The direction of motion of both stimuli was always towards the fixation point to aid fixation. Each psychometric function was based on the results of 64 trials. The point of subjective equality was taken as a measure of the perceived speed of the modulation. Each data point in the figures was based on the mean of the results from at least two such runs. The authors served as subjects and data collection preceded the modelling phase.

RESULTS

Experiment 1

The motion of a contrast modulation of a sine-wave grating is a member of a class of non-rigid motion signals which have been studied extensively in the context of the theory of wave propagation. More recently, Fleet and Langley (1994) have applied this theoretical perspective to the study of a range of non-Fourier motion patterns. For non-rigid motion patterns the orientation of iso-brightness contours in the space–time representation varies over space and time and therefore the local velocity field is spatially structured. If we can recover the local velocity field accurately we should expect perceived speed to vary in different regions of the display. However for a moving contrast modulation the motion signal defining the motion of the envelope is everywhere equal and constant. If our motion percept reflected the motion of the envelope *per se*, perceived speed should be independent of the spatial position or the motion of the carrier.

We measured perceived speed of the high contrast and low contrast regions of the contrast-modulated gratings at three carrier speeds. The speed of the modulation was always 1 deg/sec. The spatial frequency of the modulation was 1.5 c/deg and the carrier 6.0 c/deg. The speed of the carrier was either 0.25, 0.0 or −0.25 deg/sec. The data are shown in Fig. 3. In the low contrast regions perceived speed followed the direction and speed of the modulation although speed was underestimated. In the high contrast regions perceived speed followed the direction and speed of the carrier, although both subjects reported some induced motion in the reverse direction to the motion of the modulation for static carriers. This effect was not pursued further in the present series of experiments. We found that subjects could make judgements about the motion of localized regions of the display. The fact that, in the case of the opposing carrier and envelope motion, perceived motion was approximately equal and opposite, demonstrates that subjects were not simply averaging motion signals over space. For high envelope speeds or high modulation spatial frequencies the perception of the smooth motion of the low contrast regions of the display started to break up into a much more complex chaotic, "flickering" motion percept. In this study the parameters of the modulation were limited to the region in which the subjects had a clear percept of smooth motion of the low contrast regions of the signal.

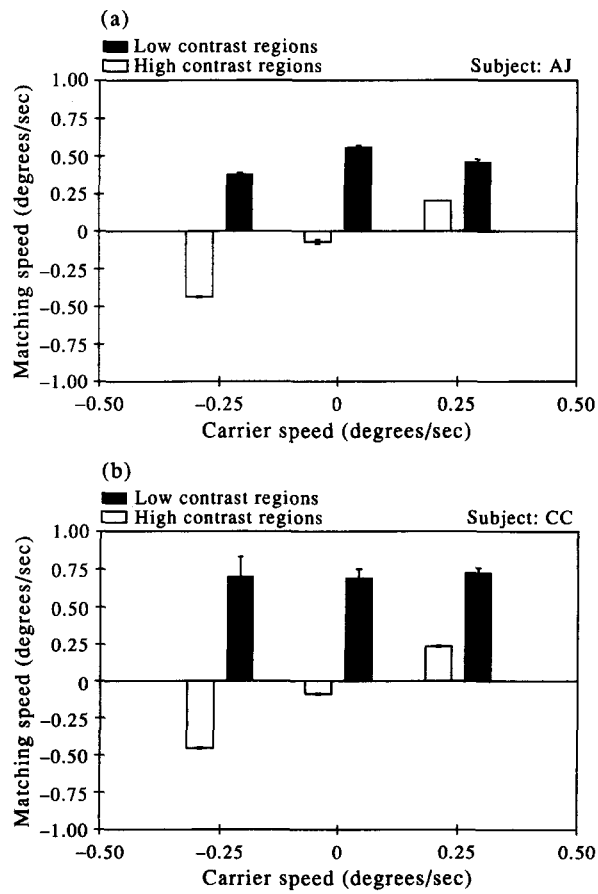

FIGURE 3. Perceived speed of low and high contrast regions of the contrast-modulated sine grating for carrier speeds of 0.25, 0.0 and −0.25 deg/sec. Modulation spatial frequency = 1.5 c/deg; carrier spatial frequency = 6.0 c/deg. (a) Subject AJ. (b) Subject CC.

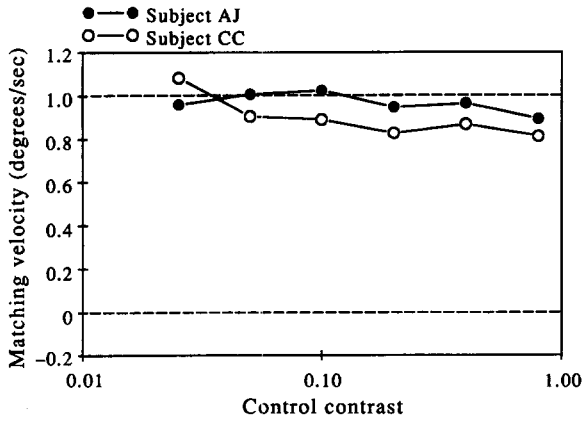

FIGURE 4. The grating speed necessary to match the perceived velocity of the low contrast regions of contrast-modulated gratings plotted as a function of the contrast of the sine grating. The contrast-modulated grating had a modulation frequency and carrier frequency of 1.5 c/deg.

Experiment 2

There is considerable evidence that the lower threshold for motion (Harris, 1984; Johnston & Wright, 1985), perceived speed (Campbell & Maffei, 1981) and speed discrimination (McKee, Silverman & Nakayama, 1986; Turano & Pantle, 1989) are unaffected by contrast for contrasts above 0.05. However there have been some recent reports that contrast can have an effect on perceived speed at much higher contrasts (Stone & Thompson, 1992). We matched the contrast of the sine gratings to the mean contrast of the contrast-modulated patterns for a modulation depth of 1.0. By definition, there is no method of matching contrast locally. Note that in the contrast-modulated patterns with static carriers perceived motion is greatest in regions where perceived contrast is low. Nevertheless, we wished to ensure that the choice of 25% contrast for the sine gratings did not have a radical effect on the measurement of the speed of the contrast-modulated patterns. We measured the speed of the low contrast regions of a contrast-modulated pattern using a range of contrasts from 0.025 to 0.8 in logarithmic steps. For the contrast-modulated pattern the modulation and carrier frequency was set to 1.5 c/deg and the speed of the modulation was 1 deg/sec. In our experimental paradigm we find little effect of contrast (Fig. 4). There is a slight slowing of the test grating as contrast is reduced but no marked effect over the range of contrasts tested here.

There are a number of differences between experimental conditions used here and those of Stone and Thompson. Their findings may depend upon the short durations used or the shaping of the stimulus induced by the use of Gaussian spatio-temporal windows. Müller and Greenlee (1994) recently confirmed that the lower threshold for motion and velocity discrimination thresholds were unaffected by contrast for contrasts above 1%, although they also report an effect of contrast on perceived speed for contrasts up to 30%. Adaptation was found to introduce or increase an effect of contrast on speed perception.

Experiment 3

We measured the perceived speed of the low contrast regions of contrast-modulated patterns as a function of the spatial frequency of the carrier grating. In all conditions the modulation velocity was set to 1 deg/sec. The spatial frequency of the modulation and the spatial

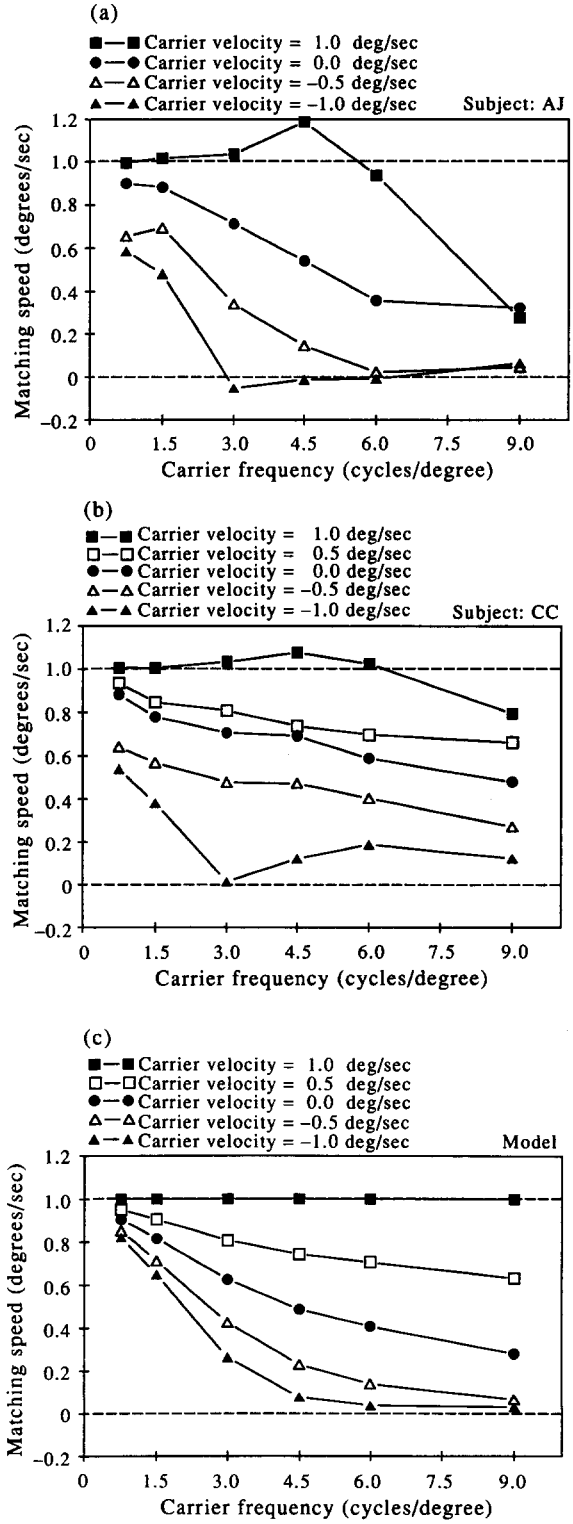

FIGURE 5. The grating speed necessary to match the perceived velocity of the low contrast regions of contrast-modulated gratings plotted as a function of the carrier frequency and velocity. The spatial frequency of the modulation and of the sine grating is 1.5c/deg. (a) Subject AJ. (b) Subject CC. (c) The results of the model.

frequency of the luminance grating was set to 1.5 c/deg. If motion perception relied on an ideal mechanism for the recovery of envelope motion in which a full-wave rectification stage was followed by subsequent filtering and motion energy analysis, we would predict that the motion field would be uniform and that the motion of the contrast modulation would appear to match the motion of a luminance grating moving at the same speed irrespective of carrier frequency.

The data are presented in Fig. 5. We found that for a static carrier the perceived speed of motion declined as the spatial frequency of the carrier increased. We also investigated a condition, "theta motion", described by Zanker (1993). In theta motion first- and second-order motion signals are opposed. When the carrier moved in the opposite direction to the contrast modulation, we found perceived motion was reduced and in fact the motion of the low contrast regions either could not be perceived or was profoundly slowed for carrier frequencies above 3 c/deg. When the motion of the carrier was in the same direction and had the same speed as the carrier, motion perception was veridical, as might be expected, since in this case all the energy in the frequency domain lies on a diagonal line through the origin, although one of the subjects reported slowing at high carrier spatial frequencies. For a carrier velocity of 0.5 deg/sec the magnitude of the reduction in perceived velocity lay between the 1 deg/sec and static carrier conditions. For static carriers we noticed that in each case the low contrast regions of the contrast-modulated patterns appeared to move faster than other parts of the display, however this effect was most clear in the case of the high spatial frequency carriers. For static carriers on some occasions it was possible to see induced motion in the reversed direction but these effects were not investigated in any detail.

Perceived velocity of the low contrast regions of contrast-modulated gratings depended upon the speed of the carrier. One conclusion that can be drawn from this is that perceived speed in contrast-modulated patterns does not depend purely upon the local orientation of energy in the frequency domain, the group velocity (Fleet & Langley, 1994), which is unaffected by this manipulation. The second is that performance is not dependent upon the operation of an ideal envelope detector based on the operation of a full-wave rectifying pre-filter because, irrespective of the carrier speed, full-wave rectification provides a first-order signal in which the dominant Fourier components lie on a line through the origin, the orientation of which is determined by the speed of the contrast modulation.

We derived predictions for the speed matching task using the McGM v.3 described in Johnston and Clifford (1995). The filter operations included in the model and the values of the model parameters were chosen to simulate a number of effects found in studies of apparent motion. These included the reversed apparent motion illusion and its elimination in the presence of a grey inter-frame interval (IFI) (Shioiri & Cavanagh, 1990), the missing fundamental illusion and its elimination in the presence of a grey IFI (Georgeson & Harris, 1990) and the induction of reversed motion in an apparent motion sequence from the introduction of a grey IFI (Pantle & Turano, 1992). Simulation of these effects allowed us to place considerable constraints on the parameters of the model and the parameters were left unchanged for the current simulations.

When presented with a translating pattern the response of the model is phase invariant and all the units in the output signal the correct speed. The values of the matching speed ascribed to the model when presented with a contrast-modulated grating were calculated by taking a space–time average of the responses to motion in the direction of the drift of the modulation. The averaging procedure was as follows. The low contrast region of a sinusoidal modulation was defined as the quarter cycle around the minimum of the envelope. This region appeared to encompass the main source of signals on which the subject based his response. The response to motion in the direction of drift of units falling within these regions was summed over space and time. To give a matching speed this sum was divided by the area of the region in the space–time image.

In order to provide a means of translating between the space–time units of the modelling space and those of the real world, we measured the high frequency limits of the envelope of our spatio-temporal differentiating filters. We assume a high spatial frequency limit of 45 c/deg for 1.5 deg visual eccentricity, which was calculated using the spatial scaling equations for peripheral vision described by Johnston (1987, 1989), with an assumption of 60 c/deg limit for foveal vision and an α parameter of 9.0 deg. We took the critical flicker fusion limit to be 60 Hz. In the time domain, 1 c/image is equivalent to 1 c/sec with each pixel representing a frame duration 7.8 msec. To estimate the amount of visual space represented by 1 pixel we assumed differentiating filters up to order 8 and matched the upper spatial frequency limit with that of the human system. This procedure gave a value of 44 sec arc per pixel making 1 c/image equivalent to 0.64 c/deg. Details of the filters and filter parameters are given in the Appendix of Johnston and Clifford (1995). Thus the only change in the model from that used in the earlier study was a re-scaling brought about by the need to accommodate the effect of viewing the stimuli in the near periphery. This re-scaling was necessary to provide a good fit to the data.

The results of the model [Fig. 5(c)] follow the main features of the experimental data. There is a reduction in computed speed for static carriers and computed speed falls to zero for theta motion when the carrier spatial frequency exceeds around 3 c/deg. When the carrier has the same direction and speed as the contrast modulation, speed is computed veridically.

In Fig. 5(c) we set the region of interest to 0.25 cycles based on our judgement of the extent of the low contrast regions after inspection of the contrast-modulated pattern. To indicate the dependence of the modelled results on this parameter, we plot the predictions for perceived speed (1.5 c/deg envelope; 6 c/deg carrier) as a function of the carrier velocity in Fig. 6(a), with dotted

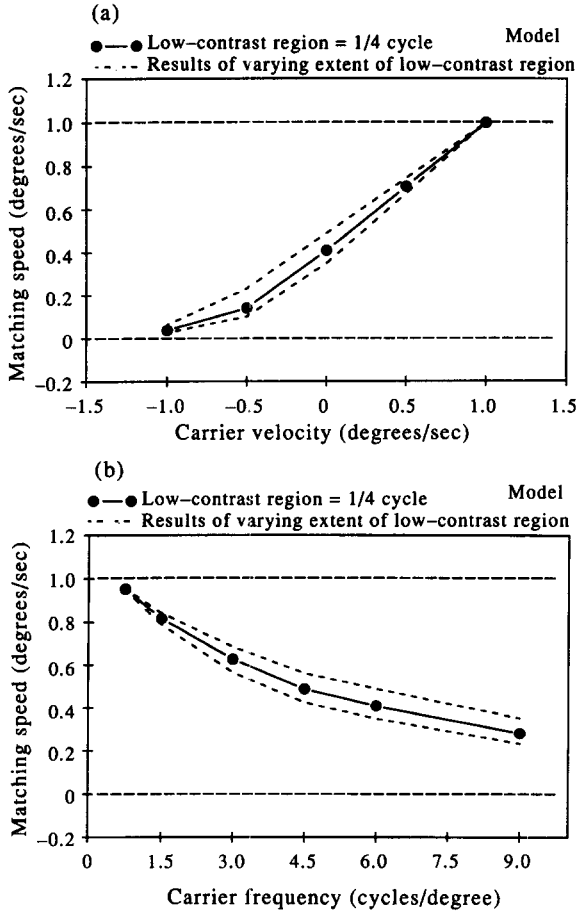

FIGURE 6. The effect of varying the region of interest parameter used in the analysis of the results of the model. (a) Model predictions of perceived speed as a function of carrier speed for a 6 c/deg carrier frequency. The solid line is for a region of interest of 0.25 cycles of the modulation signal. The upper dotted line is for 0.15 cycles; the lower dotted line for 0.35 cycles. (b) Model predictions of perceived speed as a function of carrier frequency for a static carrier. The solid line is for a region of interest of 0.25 cycles of the modulation signal. The upper dotted line is for 0.15 cycles; the lower dotted line for 0.35 cycles.

lines showing the results for window sizes of 0.15 and 0.35 cycles. The choice of window size is not critical for the modelled results. The greatest variation is seen for the static carrier. Figure 6(b) shows the variation in the modelled results for the static carrier condition as a function of carrier frequency using the same scheme.

Experiment 4

The stimuli used in Expt 3 are fairly well localized in frequency space since they can be constructed from three sine-wave components, and in the image domain the brightness varies smoothly, although we should point out that they are presented within windows with sharp spatial boundaries which will lead to a spreading of energy in the Fourier domain. One might think that these factors would favour the gradient approach since within the bounds of the stimulus there are no sharp discontinuities in the intensity surface. Similarly, an increase in the number of components along the line in frequency space might aid recovery of the orientation of energy in the Fourier domain. In Expt 4 we measured the perceived speed of motion of the low contrast regions in amplitude-modulated gratings in which the amplitude modulation had a square-wave profile. The frequency space description of this stimulus has oriented energy which contains the odd harmonics of the fundamental frequency of the square wave. The location of these bands of energy is determined by the carrier frequency [Fig. 1(i, j)].

Figure 7 shows data for two subjects and the predictions from the model. For a square-wave contrast modulation the low contrast region must be defined differently. In this case the luminance of the body of the low contrast region is uniformly grey, and it is only the regions where contrast changes abruptly that contain motion information [Fig. 1(k)]. For this reason the regions of interest used to model the perceived motion of square-wave envelopes were taken to be those regions which overlapped the contrast discontinuity. The width of each region of interest was one-eighth of the cycle. So, again one-quarter of the image was designated "low contrast". As before, the perceived speed of motion of the modulation declined as a function of the speed of the carrier.

Experiment 5

We measured perceived speed as a function of modulation depth for amplitude-modulated gratings. Turano and Pantle (1989) found velocity discrimination thresholds increased at low modulation depths. A modulation depth of 1.0 indicated a change in contrast between twice the average value and zero. A modulation depth of zero provides a sine grating at the carrier frequency. Changing the amplitude of the modulation affects the energy in the Fourier transform of the contrast-modulated gratings without changing the location of the components or their relative orientations. Therefore a strict application of the MFFC principle based on full-wave rectification, and selective filtering followed by motion energy analysis and recovery of the orientation of the energy through the origin in frequency space, would not predict changes in perceived speed

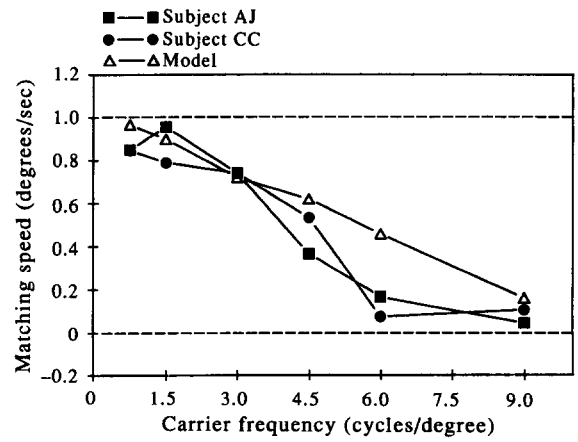

FIGURE 7. The grating speed necessary to match the perceived velocity of the low contrast regions of contrast-modulated gratings plotted as a function of the carrier velocity for square-wave modulation. The spatial frequency of the square-wave modulation and of the sine grating is 1.5 c/deg.

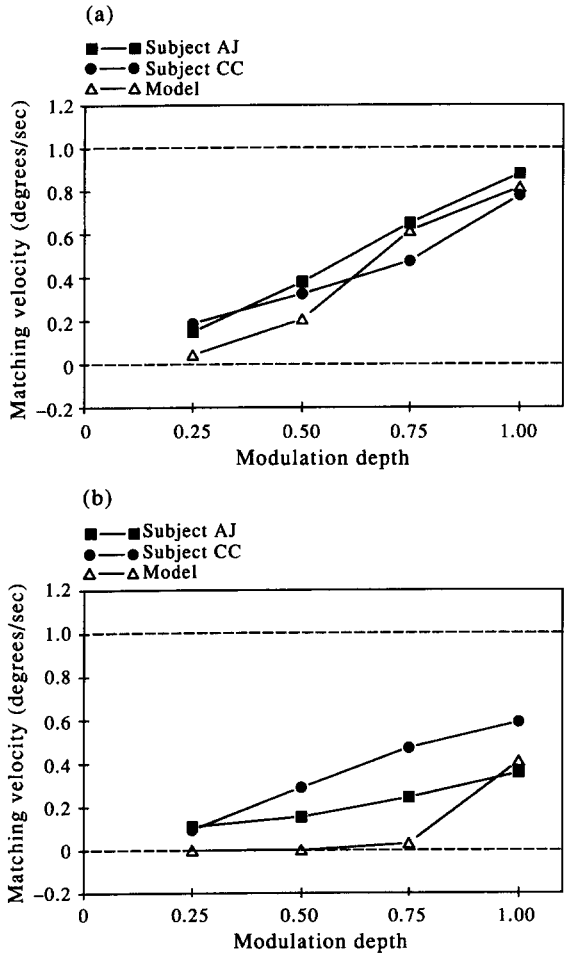

FIGURE 8. The grating speed necessary to match the perceived velocity of the low contrast regions of contrast-modulated gratings plotted as a function of the modulation depth. The spatial frequency of the modulation and of the sine grating is 1.5 c/deg. In (a) the carrier spatial frequency is 1.5 c/deg and in (b) 6.0 c/deg.

due to changes in modulation depth. Similarly, the manipulation does not affect the group velocity, the local orientation of energy in the Fourier domain. Data were collected for two carrier spatial frequencies, 1.5 and 6.0 c/deg, and are plotted as a function of modulation depth (Fig. 8). We found that reducing modulation depth reduced perceived velocity for both carrier frequencies. As expected, perceived speed was lower for the 6.0 c/deg carrier than for the 1.5 c/deg carrier. The model tended to underestimate the perceived speed of the modulation for the 6.0 c/deg carrier at low modulation depths but perceived speed is already a very small percentage of the modulation speed for these motion patterns.

Experiment 6

In order to investigate whether the reductions in speed depended upon the relative magnitudes of the envelope and carrier frequencies, we measured the perceived speed of the low contrast regions as a function of carrier spatial frequency at an additional modulation frequency of 0.75 c/deg for static carriers. The data are plotted in Fig. 9(a, b) along with the data for static carriers for a modulation frequency of 1.5 c/deg replotted from Fig. 5(a, b). The data are plotted against the log of the spatial frequency of the carrier. If perceived speed depends upon the relative magnitudes of the envelope and carrier frequency, the data should fall on two parallel lines. There were no systematic differences in the psychophysical data for the two modulation frequencies tested. There is a measurable difference in the predictions of the model for 1.5 and 0.75 c/deg [Fig. 9(c)] and the differences are reasonably consistent with the principle that the matching speed depends upon the relative frequencies of modulation and carrier, at least for carriers that are high multiples of the modulation frequency. However the predicted differences are not large and any psychophysical effect may have been swamped by random variation in the data.

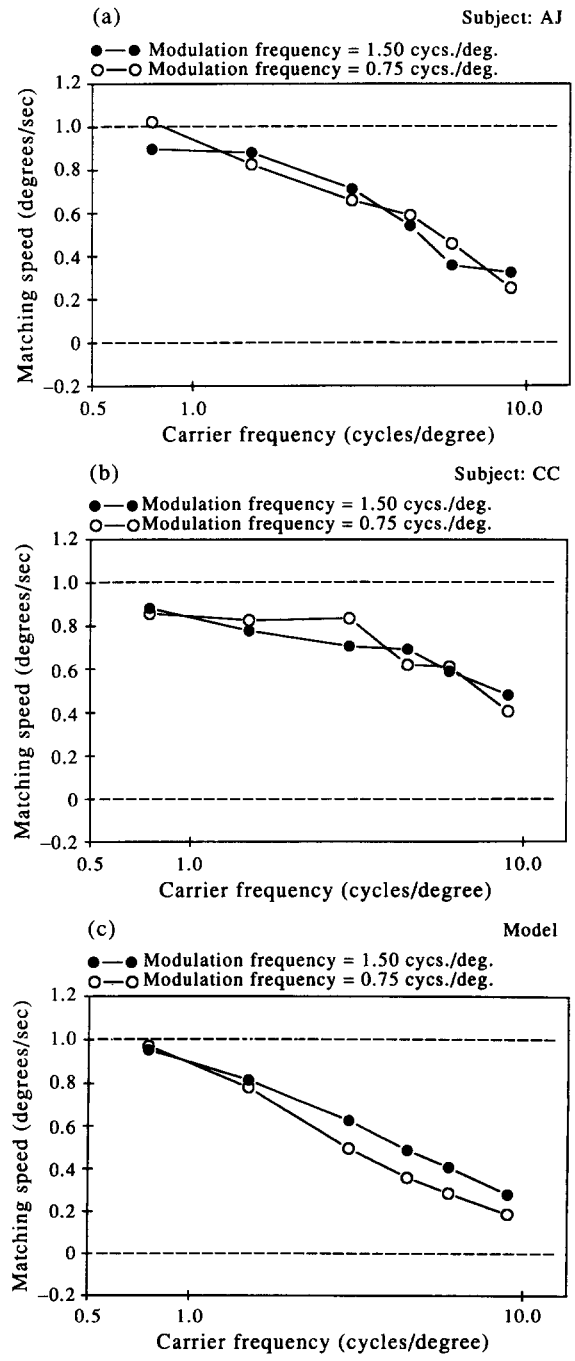

FIGURE 9. The grating speed necessary to match the perceived velocity of the low contrast regions of contrast-modulated gratings plotted as a function of the carrier frequency for two modulation frequencies 1.5 and 0.75 c/deg. (a) Subject AJ. (b) Subject CC. (c) Model predictions.

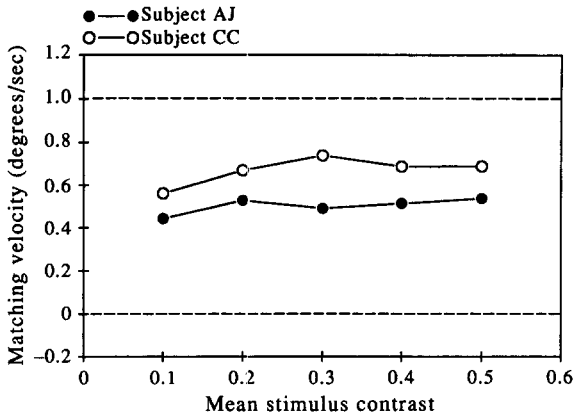

FIGURE 10. The grating speed necessary to match the perceived velocity of the low contrast regions of contrast-modulated gratings plotted as a function of the carrier contrast.

Experiment 7

It is possible that the reductions in perceived speed as a function of carrier spatial frequency described in Expt 4 might be due to a reduction in the sensitivity to the carrier with increased spatial frequency. This is unlikely given that the range of spatial frequencies used here ran from 0.75 to 9 c/deg and increasing the speed of low spatial frequency carriers, which would be expected to aid visibility (Robson, 1966), reduced the perceived speed of the low contrast regions. However, in order to investigate whether the reduction in perceived speed may have been due to a reduction in the strength of the carrier signal, we measured the perceived speed of the low contrast regions for a 1.5 c/deg modulation and a 6 c/deg static carrier as a function of the contrast of the carrier. The test grating had a spatial frequency of 1.5 c/deg and a contrast of 0.25. The data are presented in Fig. 10. There is little effect of contrast for average contrasts over 0.1. We find the speed of the low contrast regions was underestimated even in the case of high contrast carriers.

Turano and Pantle (1989) found that velocity discrimination thresholds increased as the modulation depth decreased but there was little effect of contrast. We see the same pattern of results here in the case of perceived speed. Reducing mean contrast reduces the contrast of each of the three sine-wave components making up the signal, reducing the modulation depth reduces the contrast of the sidebands relative to the carrier. We would support their conclusion that these findings demonstrate that subjects are not simply utilizing the signal in the sidebands and add that neither perceived speed nor speed discrimination depends upon the energy in the sidebands *per se*.

DISCUSSION

Do we have a special mechanism for the analysis of second-order motions?

Two recent models of motion perception propose multiple motion subsystems in order to explain the perception of motion in second-order motion stimuli (Chubb & Sperling, 1988, 1989; Kim & Wilson, 1993; Wilson *et al.*, 1992). One subsystem analyses image motion on the basis of the MFFC principle, which involves measurement of the polar angle of localized energy in the frequency domain, and another applies motion energy analysis after full-wave rectification of a band-pass filtered signal (Fig. 2). In the Wilson, Fererra and Yo model, motion energy analysis is performed at a lower spatial scale in the non-Fourier system. This would have the effect of selecting components at the modulation frequency. They also include an explicit integration stage. We can make some clear predictions in the case of contrast-modulated motion patterns processed by a non-Fourier pathway, the pathway most appropriate for contrast domain stimuli, assuming this route acts as an ideal envelope detector, (i) the motion field should be spatially unstructured, (ii) perceived motion should be veridical and (iii) the carrier frequency should have no effect on perceived motion. We found that for static carriers perceived speed was greatest in the low contrast regions of the display. In general the speed of motion of the low contrast regions was underestimated and in some cases no motion could be discerned. The perceived motion in the low contrast regions depended upon carrier frequency and carrier speed. Thus a single ideal non-Fourier channel could not account for the data presented above.

Full-wave rectification of a contrast-modulated pattern has the effect of shifting the most prominent components of the motion signal to lie on a line through the origin in Fourier space thus making the signal amenable to motion energy analysis. Alternatively, we might try to recover the local orientation of energy in Fourier space, the group velocity (Fleet & Langley, 1994). Again, ideal recovery of the group velocity would predict a veridical and spatially uniform motion percept, a prediction which was not found to be supported by the data. Zanker (1993) has suggested that we could recover theta motion by utilizing a two-stage process in which the output of spatio-temporal correlators are fed to an additional stage of motion analysis. This, in effect, transforms a second-order stimulus into a first-order stimulus and would lead to better performance than is seen in the psychophysical data.

More recently, McGowan and Chubb (1994) have argued for an additional non-Fourier motion mechanism with a broadly tuned spatial pre-filter most sensitive at high spatial frequencies, followed by full-wave rectification and motion analysis. Solomon and Sperling (1994) have also proposed an additional non-Fourier mechanism, in their case a channel in which motion energy analysis is applied after half-wave rectification. Both models include the full-wave rectification pathway, which is the main focus of the current paper, as well as the Fourier route.

Since Braddick's (1974) distinction between short-range and long-range motion there have been many further proposals of multiple motion mechanisms, often to explain isolated phenomena, which if factorially combined would lead to a combinatorial explosion of separate subsystems. The alternative, a single mechanism

which responds in different ways depending upon stimulus parameters, is rarely considered but Johnston and Clifford (1995) demonstrate that a number of apparent motion paradigms that have led to proposals of multiple mechanisms can be accommodated by a single model. Multiple systems providing multiple answers to the question of recovering local speed lead to problems of integration. We see this as a dilemma which should lead us to re-examine the evidence for multiple mechanisms rather than a problem to which a solution may be found. Solomon and Sperling (1994) explicitly discount the need for integration, favouring parallel access, leading to a three-way transparency. The problem for this approach is to explain the coherence of the visual percept under most circumstances.

We were unable to find experimental support for the idea that there exists in the human system a special mechanism for the accurate detection of the motion of the envelope of an amplitude-modulated signal based on rectification followed by further filtering and motion energy analysis. We now consider whether perceived motion in rigid translations and contrast-modulated motion patterns can be accounted for by a single motion mechanism.

The McGM v.3

The McGM (Johnston & Clifford, 1995; Johnston *et al.*, 1992) offers a rapprochement between evidence for multiple spatio-temporal channels in the human and primate visual system and gradient schemes for calculating image speed which are prominent in computer vision. As Smith and Edgar (1994) point out, the current view is that there are two or at most three temporal filters spanning the visible range although the linear spatial filters are more numerous and narrow-band. We have reported in an earlier paper that three temporal filters are necessary to simulate reversals of motion in some sampled motion illusions. In the McGM the filters are derivatives of Gaussians in space and derivatives of log Gaussians in time (see Koenderink, 1988). Derivatives of Gaussians have been shown to provide good models of cortical cell receptive fields (Young & Lesperance, 1993) and we have demonstrated (Johnston & Clifford, 1995) that the temporal filters used in the model provide close fits to psychophysical data on the number and shape of the temporal filters in the human visual system (Hess & Snowden, 1992).

In gradient schemes speed is recovered by forming a quotient between measures of the change of image intensity with respect to time and space. The advantage of forming a quotient from a biological perspective is that it allows a biological system to become insensitive to attributes of the image, like contrast or chromaticity, which are irrelevant to the computation at hand. If two neurones are equally sensitive to contrast and wavelength but differentially sensitive to spatio-temporal frequency, a measure of the relative firing of the two units will be insensitive to contrast and wavelength, but the unit would carry information about spatio-temporal change (Johnston *et al.*, 1992; Johnston & Wright, 1986). Note that if a neurone codes information about speed, we might expect a saturated response to contrast on the basis of stimulus selectivity without regard to the existence of an explicit gain control mechanism.

The idea that temporal frequency might be encoded by comparing the relative response of filters with different temporal frequency tuning curves has been supported by a number of authors (Harris, 1980; Smith & Edgar, 1994; Thompson, 1982; Tolhurst, Sharpe & Hart, 1973). Harris (1980) showed that the ratio of pattern to flicker thresholds depended upon velocity irrespective of the spatio-temporal description of the sine-wave gratings used to measure the thresholds, and recently Smith and Edgar (1994) provide evidence of *increases* in perceived speed for gratings at high temporal frequencies after adaptation to the motion of gratings of low temporal frequencies. These findings provide strong support for the idea that speed is computed by means of a comparison between filters which vary in their temporal frequency tuning. However, if just temporal frequency is recovered in this way, a measure of the spatial frequency of the pattern would also need to be obtained to recover image speed. Johnston and Wright (1985, 1986) pointed out that if the filters differed in their spatial tuning as well as their temporal tuning characteristics, image speed could be recovered directly from the comparison stage. Although it is straightforward to consider the relative activity in two separate spatio-temporal mechanisms, there is no obvious and immediate generalization of this approach to deal with multiple spatio-temporal filters. However, we know that we have three different temporal filters in the human visual system, and an as yet undetermined number of spatial filters.

The McGM describes how filters sensitive to a range of spatial frequencies can be combined. The current version of the model operates at a single spatial scale set by the parameters of the zero-order Gaussian kernel. However, derivative of Gaussian filters become tuned to higher spatial frequencies as the order of differentiation increases, although they are sensitive to spatial frequency across the range, and thus the model would predict a range of independent spatial channels at threshold. It would of course be possible to incorporate additional filter combinations based on blurring kernels at different scales if necessary. We also incorporate three temporal filters. The central computational strategy involves a comparison between differentiating filters of the same order, one of which substitutes a temporal differentiation for a spatial differentiation. Pairs of filters are combined using a weighted least-squares procedure since a quotient formed from any pair taken in isolation would be ill-conditioned.

The least-squares algorithm requires the computation of a sum of squared terms, which forms the denominator of the quotient. Each term therefore involves a squaring non-linearity. This aspect of the model draws support from Heeger (1992) who provides strong evidence for the presence of squaring non-linearities in cat cortical cells. He also incorporates a threshold non-linearity since single neurones cannot represent a change in sign. Note that in

the McGM rectification is a prediction of the model which emerges from the algorithm rather than a component part or feature of the design specification of the model. The role of non-linearities and rectification in the McGM is one of conditioning the quotient to derive a robust measure of image speed from spatio-temporal gradients of image brightness.

Signal decomposition or surface geometry?

We have juxtaposed two quite different approaches to the understanding of image motion analysis. In the MFFC approach the emphasis is placed on signal decomposition and how spatio-temporal filters may be organized to select components from a complex signal. Henning, Hertz and Broadbent (1975) adequately demonstrated that the visual system does not have independent access to individual components of static amplitude-modulated gratings. Nevertheless, it is clear that for the class of signals considered here it is quite possible to envisage a pipeline of filtering operations which would allow the correct recovery of the speed of the envelope in contrast-modulated patterns and a number of models are considered above. Since we have shown that perceived motion of the envelope is not veridical, to support the signal decomposition approach, we would have to argue that the pattern of data reflect systematic error in the calculation of the envelope speed.

The alternative approach is to place the emphasis on

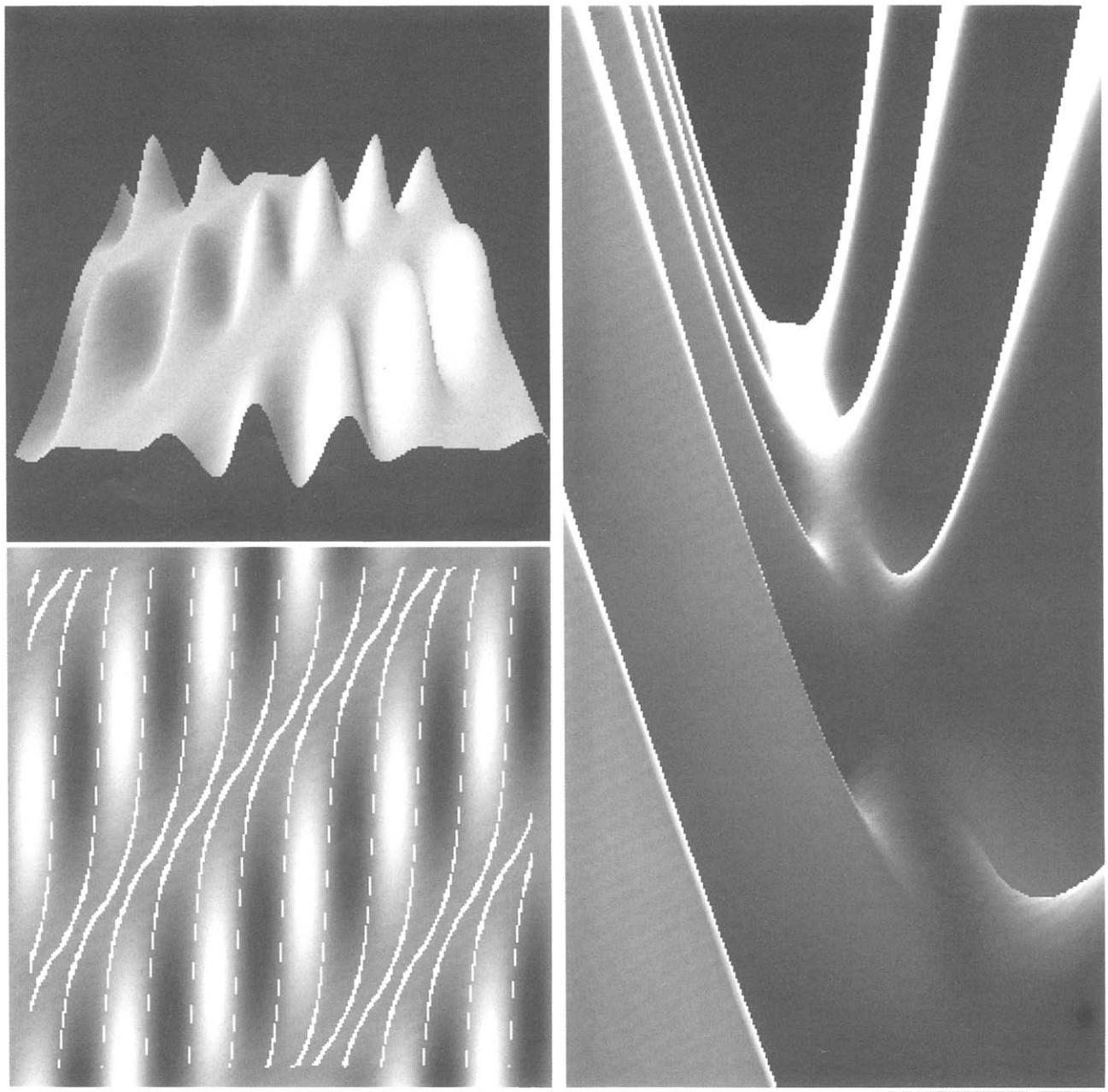

FIGURE 11. Top left: a shaded surface in which height represents image brightness as a function of space and time for a contrast-modulated grating. Modulation frequency = 1.5 c/image; carrier frequency = 6 c/image. Right: an expanded version of the top left showing a detail of the surface in the low contrast regions. Bottom left: a contrast-modulated grating stimulus with the zero-crossings of the Laplacian derivative superimposed to indicate the orientation of points of inflection in the brightness surface.

the geometry of the surface defined by the brightness values in the space–time image (Fig. 11). In this context the McGM simply provides a robust means of calculating the local velocity field for a variety of stimulus conditions using biologically plausible filtering operations. In the spatio-temporal gradients approach the spatio-temporal filters in the visual system act as blurred differentiating filters which calculate geometric properties of the blurred surface Koenderink & van Doorn, 1987). The most intuitive way to derive predictions from a surface gradient model is to consider how a ball or a raindrop would run off the surface. In Fig. 11 (top left) the brightness surface of the contrast-modulated signal is defined by shading. For a grating, a cylindrical surface, the ball would always run off in the same direction at any point, the direction orthogonal to the iso-brightness contours of the image. The orientation of the iso-brightness contours gives the speed. For the contrast-modulated pattern it is clear that the orientation of the gradient varies from point to point. In the high contrast areas the iso-brightness contours are oriented vertically since they are dominated by the carrier signal. In the low contrast regions the surface is influenced to a greater extent by the envelope, shaping the surface, and a ball would run into or away from the plateaux in the low contrast region. In Fig. 11 (right) we provide an expanded version of the surface in the low contrast regions viewed from a angle in which the line of sight is aligned with the envelope to show the local structure of the surface in the low contrast regions of the stimulus.

Figure 11 (bottom left) shows the original image with the zero-crossings of the Laplacian derivative (Marr & Hildreth, 1980) superimposed. The zero-crossing map marks points of inflection on the surface. The points of inflection separate maxima and minima on the surface and therefore give an indication of the orientation of the gradient, although we should point out that the direction of the gradient is not necessarily perpendicular to the zero-crossings. In the low contrast regions the orientation of the zero-crossing contour is shifted towards the vertical relative to the envelope; thus there is oriented structure in the image which could give rise to the slowing predicted by the model and seen in the data. It can also be seen from Fig. 11 (bottom left) that an increase in the spatial frequency of the carrier would shift the orientation of the zero-crossings further towards the vertical. Again we can interpret this as indicating a slowing, in this case with an increase in the spatial frequency of the carrier; the result shown in Fig. 5. The model, of course, gives a quantitative measure of the orientation of the gradient in the conditions discussed informally here. Thus we can see that the changes in perceived speed in the low contrast regions of amplitude-modulated patterns follow naturally from an analysis of the local brightness gradients in the space–time image and reflect the computed values of the local velocity field.

CONCLUSIONS

In the Chubb and Sperling (1988, 1989) and Wilson *et al.* (1992) models full-wave rectification is a theoretical construct, an explicit pre-processing stage which forms part of a special sub-system for the accurate measurement of second-order motion. However, we have shown that the accurate recovery of the motion of a contrast modulation is a task beyond the normal capacities of the human motion analysis system. In the McGM, a general purpose motion analysis scheme, rectification is a consequence of the mechanism used to condition the quotient for the recovery of image speed. The model provides accurate quantitative predictions for perceived motion in contrast-modulated motion displays.

REFERENCES

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2*, 284–299.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research, 14*, 519–527.

Campbell, F. W. & Maffei, L. (1981). The influence of spatial frequency and contrast on the perception of moving patterns. *Vision Research, 21*, 713–721.

Chubb, C. & Sperling, G. (1988). Drift-balanced random dot stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5*, 1986–2007.

Chubb, C. & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Science U.S.A., 86*, 2985–2989.

Derrington, A. M. (1987). Distortion products in geniculate X cells: A physiological basis for masking by spatially modulated gratings. *Vision Research, 27*, 1377–1386.

Derrington, A. M. (1990). Mechanisms for coding luminance patterns: Are they really linear? In Blakemore, C. (Ed.), *Vision coding and efficiency* (pp. 175–184). Cambridge: Cambridge University Press.

Derrington, A. M. & Badcock, D. R. (1985). Seperate detectors for simple and complex grating patterns? *Vision Research, 25*, 1869–1878.

Derrington, A. M., Badcock, D. R. & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research, 33*, 1785–1794.

Fennema, C. L. & Thompson, W. B. (1979). Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing, 9*, 301–315.

Fleet, D. J. & Langley, K. (1994). Computational analysis of non-Fourier motion. *Vision Research, 34*, 3057–3079.

Georgeson, M. A. & Harris, M. G. (1990). The temporal range of motion sensing and motion perception. *Vision Research, 30*, 615–619.

Harris, M. G. (1980). Velocity specificity of the flicker to pattern sensitivity ratio in human vision. *Vision Research, 20*, 687–691.

Harris, M. G. (1984). The role of pattern and flicker mechanisms in determining the spatiotemporal limits of velocity perception—2. The lower movement threshold. *Perception, 13*, 409–416.

Heeger, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A, 4*, 1455–1471.

Heeger, D. J. (1992). Half-squaring in responses of cat striate cells. *Visual Neuroscience, 9*, 427–443.

Heeger, D. J. & Simoncelli, E. P. (1995). Model of visual motion sensing. In Harris, L. (Ed.), *Spatial vision in humans and robots*. Cambridge: Cambridge University Press. In press.

Henning, G. B., Hertz, B. G. & Broadbent, D. E. (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research, 15*, 887–897.

Hess, R. F. & Snowden, R. J. (1992). Temporal properties of human visual filters: Number, shapes and spatial covariation. *Vision Research, 32*, 47–60.

Horn, B. K. P. & Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence, 17*, 185–203.

Johnston, A. (1987). Spatial scaling of central and peripheral contrast-sensitivity functions. *Journal of the Optical Society of America A, 4*, 1583–1593.

Johnston, A. (1989). The geometry of the topographic map in striate cortex. *Vision Research, 29*, 1493–1500.

Johnston, A. & Clifford, C. W. G. (1995). A unified account of three apparent motion illusions. *Vision Research, 35*, 1109–1123.

Johnston, A. & Wright, M. J. (1985). Lower thresholds of motion for gratings as a function of eccentricity and contrast. *Vision Research, 25*, 179–185.

Johnston, A. & Wright, M. J. (1986). Matching velocity in central and peripheral vision. *Vision Research, 26*, 1099–1109.

Johnston, A., McOwan, P. W. & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proceedings of the Royal Society of London B, 250*, 297–306.

Kim, J. & Wilson, H. R. (1993). Dependence of plaid motion coherence on component grating directions. *Vision Research, 33*, 2479–2489.

Koenderink, J. J. (1988). Scale-time. *Biological Cybernetics, 58*, 159–162.

Koenderink, J. J. & van Doorn, A. J. (1987). Representation of local geometry in the visual system. *Biological Cybernetics, 55*, 367–375.

Marr, D. & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B, 207*, 187–217.

McGowan, J. W. & Chubb, C. (1994). Evidence for a second texture-defined motion channel. *Investigative Ophthalmology and Visual Science, 35*, 1407.

McKee, S., Silverman, G. & Nakayama, K. (1986). Precise velocity discrimination despite random variation in temporal frequency. *Vision Research, 26*, 609–620.

Müller, R. & Greenlee, M. W. (1994). Effects of contrast and adaptation on the perception of the direction and speed of gratings. *Vision Research, 34*, 2071–2092.

Pantle, A. & Turano, K. (1992). Visual resolution of motion ambiguity with periodic luminance- and contrast-domain stimuli. *Vision Research, 32*, 2093–2106.

Robson, J. G. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America, 56*, 1141–1142.

Shioiri, S. & Cavanagh, P. (1990). ISI produces reverse apparent motion. *Vision Research, 30*, 757–768.

Smith, A. T. & Edgar, G. K. (1994). Antagonistic comparison of temporal frequency filter outputs as a basis for speed perception. *Vision Research, 34*, 253–265.

Sobey, P. & Srinivasan, M. V. (1991). Measurement of optical flow by a generalized gradient scheme. *Journal of the Optical Society of America A, 8*, 1488–1498.

Solomon, J. A. & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Research, 34*, 2239–2258.

Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision, 4*, 183–207.

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

Thompson, P. (1982). Perceived rate of movement depends upon contrast. *Vision Research, 22*, 377–380.

Tolhurst, D. J., Sharpe, C. R. & Hart, G. (1973). The analysis of the drift rate of moving gratings. *Vision Research, 13*, 2545–2555.

Turano, K. & Pantle, A. (1989). On the mechanism that encodes the movement of contrast variations: Velocity discrimination. *Vision Research, 29*, 207–221.

Verri, A., Straforini, M. & Torre, V. (1992). Computational aspects of motion perception in natural and artificial vision systems. *Philosophical Transactions of the Royal Society of London B, 337*, 429–443.

Watt, R. J. & Andrews, D. P. (1981). APE: Adaptive Probit Estimation of psychometric functions. *Current Psychological Reviews, 1*, 205–214.

Wilson, H. R., Ferrera, V. P. & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9*, 79–97.

Young, R. A. & Lesperance, R. M. (1993). A physiological model of motion analysis for machine vision. *Technical Report General Motors Research Laboratories, GMR-7878*, 1–76.

Zanker, J. M. (1993). Theta motion: A paradoxical stimulus to explore higher order motion extraction. *Vision Research, 33*, 553–569.

Zhou, Y. & Baker, L. C. (1993). A processing stream in mammalian visual cortex neurones for non-Fourier responses. *Science, 261*, 98–101.

Acknowledgements—The research was supported by a grant from the Joint Councils Initiative in Cognitive Science/HCI and a grant from the Wellcome Trust. We would like to thank Simon Prince for help in generating Fig. 11. Simon Prince was supported by a Wellcome Trust Vacation Scholarship.