PERGAMON

Vision Research 40 (2000) 745–758

Vision Research

www.elsevier.com/locate/visres

# Spatial resolution and receptive field height of motion sensors in human vision

Mark A. Georgeson *, Nicholas E. Scott-Samuel [1]

*School of Psychology, University of Birmingham, Birmingham B15 2TT, UK*

Received 1 June 1999; received in revised form 5 November 1999

**Abstract**

We estimated the length of motion-detecting receptive fields in human vision by measuring direction discrimination for three novel stimuli. The motion sequences contained either (i) alternate frames of two differently oriented sinusoidal gratings; (ii) alternate frames of vertical grating and plaid stimuli or (iii) a vertical grating divided into horizontal strips of equal height, where alternate strips moved leftward and rightward. All three stimulus sequences had a similar appearance (of moving strips) and the task was to identify the direction of the central strip. For sequences (ii) and (iii), performance fell as the strip height decreased. Threshold height fell with increasing contrast up to about 20%, then levelled off at the critical strip height. Temporal frequency (1.9–15 Hz) had no effect on the critical strip height. We argue that the receptive field length corresponds to twice this critical height. The length estimates were strikingly short, ranging from about 0.4 cycles at 3.0 cpd to 0.1 cycles at 0.1 cpd. These lengths agree well with the estimates derived at threshold by Anderson and Burr (1991, J. Opt. Soc. Am. A8, 1330–1339), and imply that the motion-sensing filters have very broad orientation tuning. These and other results are interpreted within the framework of a Gaussian derivative model for motion filtering. The sensitivity of motion filters to a broad range of orientations suggests a simpler view of how coherent plaid motion is processed. 

*Keywords:* Psychophysics; Motion detection; Receptive fields; Direction selectivity; Human vision; Gratings; Contrast; Direction discrimination; Gaussian derivative; Spatio-temporal filtering

## 1. Introduction

In natural vision, the speed and direction of motion in the visual field vary with spatial position, and these spatial variations carry rich information about the spatial structure of the world and about the observer's own movements. In motion analysis there must be a trade-off between the requirements for robust representation of local velocity and for resolution of changes across space. Integrating information over a relatively large region should improve the reliability of velocity estimation, but at the expense of smoothing out local variations that may be important (Braddick, 1993). While vision may handle this problem at several levels of the motion pathway, the width and height of early, motion-selective receptive fields must impose a limit on the spatial resolution of motion signals.

In this paper we aim to reveal the spatial resolution of motion perception and the corresponding receptive field height of human motion mechanisms through psychophysical experiments. We make use of an *interleaved apparent motion sequence* consisting of four frames that alternate between two types of pattern. Frame-by-frame the pattern shifts successively in one direction through 90° steps of spatial phase (Fig. 1) and with each step the pattern switches from one type to the other (symbolized by the two types of shading in Fig. 1). In general, the two types could be different colours, different kinds of modulation, different spatial frequencies and so on. The motive for using this sequence is that it consists of two stationary, counterphase flickering patterns A and B (A: 0/180° phases, B: 90/270°), interleaved in space and time (Shadlen & Carney, 1986; Georgeson & Shackleton, 1989). Thus any mechanism that is sensitive only to pattern A or only to pattern B can sense flicker, but not motion. To detect motion

* Corresponding author. Fax: +44-121-4144897.

*E-mail addresses:* m.a.georgeson@bham.ac.uk (M.A. Georgeson), ness@astra.vision.mcgill.ca (N.E. Scott-Samuel)

[1] Present address: McGill Vision Research, 687 Pine Avenue West, Montréal, Québec, Canada, H3A 1A1.

0042-6989/00/$ - see front matter 
PII: S0042-6989(99)00219-9

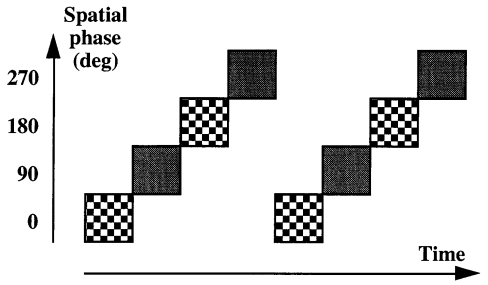


Fig. 1. The interleaved apparent motion sequence. Spatial phase steps through 90° on each frame (two cycles are shown) and the pattern alternates in some other respect (e.g. between orientations A and B) with each step. Apparent motion is possible only if there is a mechanism that can integrate the two different spatial patterns A and B symbolised by the checkered and shaded boxes, yielding motion from temporally and spatially interleaved flicker.

there must be a mechanism sensitive to both A and B. Hence the interleaved sequence can be used as a general tool for studying the pattern selectivity of motion mechanisms. It has been used already to demonstrate that motion detectors cannot combine luminance-modulated and contrast-modulated signals (Ledgeway & Smith, 1994) at least when distortion products are absent (Scott-Samuel & Georgeson, 1999).

If patterns A and B contain no common Fourier components then interleaved motion sequences are drift-balanced (Chubb & Sperling, 1988) because their spatial Fourier components are flickering, not moving. There is no net motion energy in any particular direction. However, they are not microbalanced, because viewing the images through an aperture may reveal the presence of local motions in different directions. Receptive fields (RFs) are effectively an aperture through which motion is computed, and so it may be that by asking subjects to discriminate the direction of *local* motions in these sequences we can obtain estimates of the RF sizes for motion detection. The logic of this approach is explained further below.

The experiments also address the orientation selectivity of motion mechanisms, because questions about orientation selectivity and about RF size are closely linked: a mechanism with narrow orientation tuning must have extensive lengthwise integration (a long RF) while broad orientation tuning implies a short RF. For a linear filter the RF length and orientation bandwidth are inversely related. In several experiments we examined the orientation range over which motion sensors can integrate different grating components. This was done firstly by alternating the grating orientation at each step in phase, and then by using modified sequences that progressively overcame some unexpected difficulties. van den Berg, van de Grind and van Doorn (1990) tested the orientation selectivity of motion detection with arrays of randomly oriented line segments or edges that changed orientation systematically with each motion step. Rotations of 25–40° between steps were sufficient to eliminate direction discrimination and so they concluded that the early stage of 'bilocal motion detectors' is quite narrowly tuned for orientation. As we shall see, the orientation range for motion perception can be very large — up to $\pm 80°$ — and this has the rather different implication that motion detection is a very local affair, depending on local phase shifts in the image.

In the third experiment we aim to quantify more directly the spatial extent over which motion signals are integrated in the direction orthogonal to the motion. The display contained spatially interleaved strips of sinusoidal gratings moving in opposite directions, and the idea here is that, by analogy with static spatial acuity, the narrowest strips for which opposite directions can still be resolved will indicate the scale of spatial integration. A similar logic has been used in experiments on motion segregation, with random dots drifting in opposite directions across adjacent strips (Koenderink, van Doorn & van de Grind, 1985). In that study the minimal width for resolving the separate strips of motion was about 8 min arc in foveal vision at speeds less than 1°/s, but rose markedly with both speed and eccentricity. One limitation of experiments with random dots is that the dot array is spatially broadband and so may activate a multiplicity of filters at different spatial scales under different conditions. Our experiments used moving sinusoidal waveforms in order to anchor the scale of analysis to a particular spatial frequency. Fredericksen, Verstraten and van de Grind (1997) raised a host of objections to previous methods of estimating motion RF size. Fortunately, these potential problems do not apply to our suprathreshold, acuity-based method, and it therefore serves as a valuable re-assessment of earlier estimates (e.g. Gorea, 1985; Anderson & Burr, 1987, 1991).

As noted above, RF height and orientation selectivity are linked, so it should be possible to derive values for both with the same set of stimuli. In the context of psychophysical studies, we adopt the same notion as Anderson and Burr (1991): 'The phrase psychophysical receptive field is used to denote the fact that we are measuring the psychophysical properties of perceptual units and not the physiological properties of single cells'.

## 2. Experiment 1: alternately tilted gratings

Pilot observation of a sequence of sinusoidal gratings (Fig. 2) tilted alternately to either side of the vertical and stepping 90° in phase with each presentation yielded the surprising observation that the display had

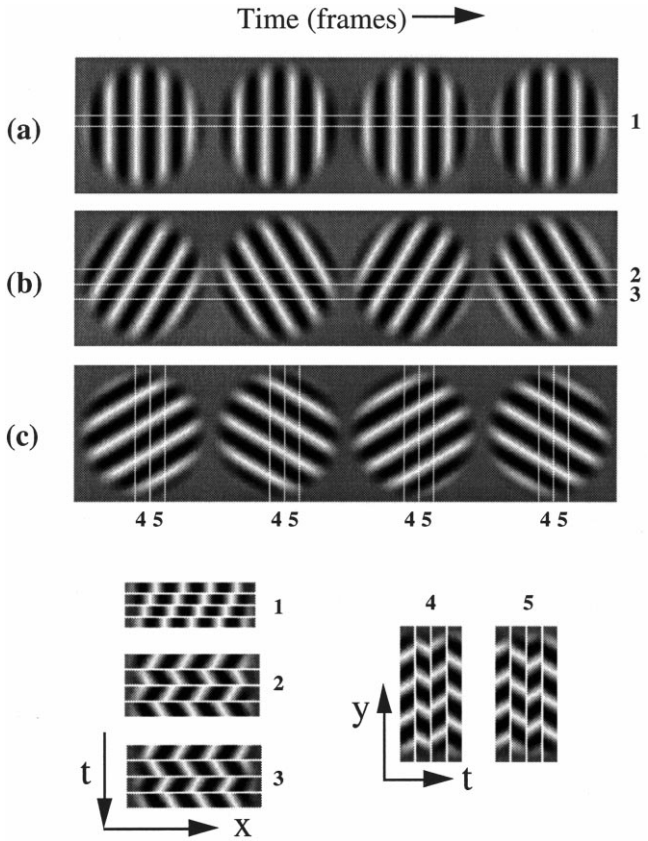


Fig. 2. Local phase changes in the tilted gratings. Experiment 1. Consecutive frames of three motion sequences (a, b, and c); spatial phase steps by 90° from frame to frame. (a) Vertical grating; (b) orientation alternates ±30° from vertical; (c) orientation alternates ±60°. The numbered strips are stacked to form space–time diagrams (below); for 1–3, time increases down the $y$-axis; for 4 and 5, time increases from left to right. Strip 1 drifts to the left; strip 2, the central strip of (b), drifts to the left while the adjacent strip 3 drifts to the right; strips 4 and 5, central and adjacent in (c), drift upwards and downwards, respectively.

the appearance of a vertical stack of horizontal strips of grating, with alternate strips moving in opposite directions (left and right). This percept changed to one of vertical strips moving up and down when the tilt of the gratings exceeded 45°. Thus consistent local motion was seen when interleaved flickering gratings were displayed at different orientations. The following experiment was designed to investigate this intriguing effect more rigorously.

### 2.1. Method

Stimuli were generated by a PC with custom-written Pascal software, and displayed via a Cambridge Research Systems VSG2/2 8-bit framestore on a gamma-corrected Eizo Flexscan 9060S monitor with a refresh rate of 60Hz. Use of two palette chips together gave the system the equivalent of 12-bit luminance resolution; that is, the full 8-bit greyscale was available even at fairly low contrasts. The mean luminance of the display was 70.5 cd/m$^2$, and its linearity was calibrated with a Minolta LS-110 photometer. Contrast was expressed as Michelson contrast, $100(L_{max}-L_{min})/(L_{max}+L_{min})$, where $L_{max}$ and $L_{min}$ were the maximum and minimum luminances, respectively.

The interleaved apparent motion sequence consisted of a sinusoidal grating whose orientation alternated between $+\theta$ and $-\theta°$ from the vertical with each 90° step in spatial phase (Fig. 2). The orientation $\theta$ was 0, 15, 30, 45, 60, 75 or 90° from vertical. The temporal frequency of the sequence was 7.5 Hz and the presentation duration was 266 ms, making each sequence two temporal cycles (eight frames) long. Contrast was fixed at 20% and the spatial frequency of the gratings was 1.0 cpd in a 4.5° field (512 × 512 pixels; 16.8 × 16.8 cm) at a viewing distance of 214 cm. On each trial there was a single presentation, and observers were required to report both the axis along which they saw motion (horizontal or vertical; the *axis task*), and the direction of motion of the strip that passed over a small, central fixation point (the *direction task*). Viewing was binocular, and the observer's head was held in position with a chin and forehead rest. The two authors acted as observers with 48 trials per observer per condition.

### 2.2. Results: motion along two axes

The results (Fig. 3, filled symbols) confirmed that for orientations $\theta \leq 45°$ strips of horizontal motion were seen, but at larger angles vertical motion was seen. Both observers reliably reported the correct direction (left/right or up/down) of the central strip, although MAG showed a dip in accuracy at 45° where the axis of motion was somewhat ambiguous. For NSS the transition between axes occurred sharply between 45 and 60°, and recognition of direction was near 100% for all angles tested.

### 2.3. Discussion: local phase shifts predict the perceived motion

An examination of the local phase changes in the experimental stimuli provides an explanation for the pattern of results. Fig. 2(a–c) shows three different sequences of the seven used, where $\theta = 0$, 30 and 60°. In sequence (a), motion is seen to the left, as space–time diagram (1) shows. In sequence (b), two adjacent strips have opposite direction (2 and 3); the local phase of the central strip (2) moves leftwards, whilst its neighbour (3) moves to the right. In sequence (c), the motion axis has now flipped to vertical, and adjacent strips move upwards (4) and downwards (5). Despite the differences in orientation of the component gratings, and the lack of moving Fourier components, these local phase changes result in reliable perception of local moving strips.

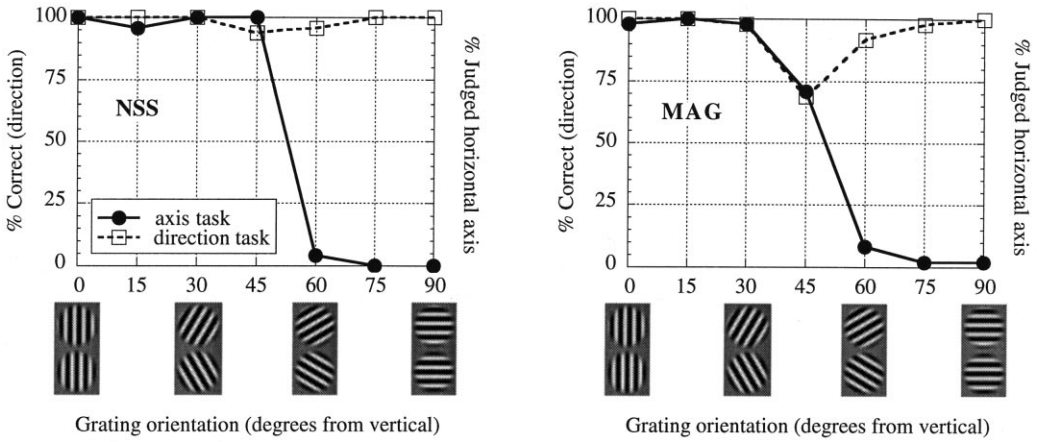


Fig. 3. Strip direction and axis results. Experiment 1. Results for two subjects (NSS and MAG) for both the axis task (filled circles, solid line) and direction task (open squares, dashed line), expressed as % correct for the direction task (left hand ordinate), and as % judged horizontal for the axis task (right hand ordinate). The $x$-axis shows the orientation $\theta$ for each sequence.

The dip in direction performance for observer MAG at $\theta = 45°$ (Fig. 3) indicates that the direction task is confounded by an uncertainty in the axis task at this value. A dip was not observed for observer NSS because there was a sharp transition in the axis task between $\theta = 45$ and 60°, and so no uncertainty was present for this observer. This transition from left/right to up/down does not mean that there is no energy along the orthogonal axis; for all values $0 < \theta < 90°$ there is both horizontal and vertical local motion present, as we show below. The transition from horizontal to vertical motion perception presumably reflects the *relative visibility* of these two motions.

A grating of spatial frequency $f$ and orientation $\theta$ has horizontal and vertical periodicities $(u, v)$ where $u = f\cos\theta$ and $v = f\sin\theta$. In the Appendix we show that if motion perception is based on local phase shifts, then the sequences used here contain horizontally moving strips whose height is $1/(4v)$, one quarter of the vertical period $(1/v)$. Informal observation confirmed that this was indeed the height of the perceived strips. Interestingly, the switching of orientations can also be described as having constant $v$ with alternation between $u$ and $-u$ (see Fig. 4) and by an entirely similar argument one can show that there are also vertical strips of upward and downward phase shift whose width is $1/(4u)$.

In principle, then, there are two alternative interpretations for each of our sequences — horizontal or vertical strips of motion. However, when $0 \le \theta < 45°$, $v < u$, and the horizontal strips are wider than the vertical ones, but when $45 < \theta \le 90°$ $v > u$ and the vertical strips are wider than the horizontal. A likely explanation for the crossover from horizontal to vertical motion perception (Fig. 3) is that two mechanisms, orthogonally tuned, respond to the two axes of motion. We assume that the two mechanisms have peak sensitivities on the $u$-axis and $v$-axis, respectively (Fig. 4). Grating components close to the $u$-axis ($\theta < 45$) will evoke a horizontal motion response more strongly than vertical, and vice versa for $\theta > 45$. Since there was little or no evidence for ambiguity or transparency in the axis task (Fig. 3) it seems likely that the stronger response suppresses the weaker — 'winner takes all'. It follows that, in spatial terms, the wider strips determine the perceived axis of motion. One might expect the crossover point to be at $\theta = 45°$, but experimentally the crossover was between 45 and 60°. It may be that extensive practice on horizontally moving stimuli created some bias in favour of the horizontal interpretation.

We did not observe oblique motion. It might seem surprising that the vertical and horizontal motions embedded in this animation do not sum to form oblique movements. Alternate rows (and columns), however, have opposite directions of movement, and so the pattern of oblique movements obtained by summation would form a checkerboard in which every small square

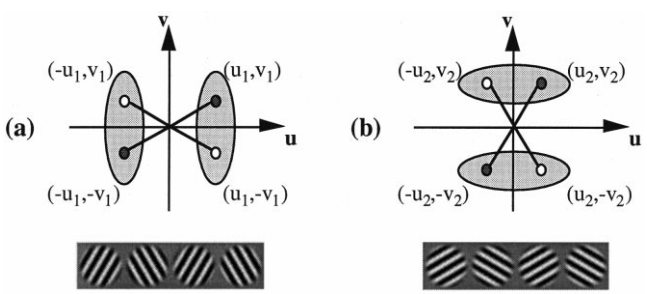


Fig. 4. Two orthogonal motion detectors. Experiment 1. (a) $\theta < 45°$ gives horizontal motion, and (b) $\theta > 45°$ gives vertical motion. The grey ovals show (a) vertically oriented detectors, (b) horizontally oriented. Note that, as $\theta$ varies, the stimuli follow a circular path through $u$–$v$ space, because grating spatial frequency was fixed $[(u_1^2 + v_1^2) = (u_2^2 + v_2^2) = \text{constant}]$.

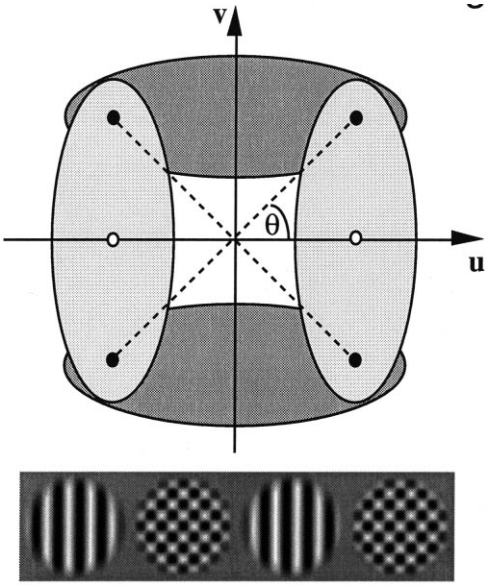


Fig. 5. The grating-plaid sequence. Experiment 2. Below: four frames of the grating-plaid motion sequence (phases 0, 90, 180, 270°; alternate frames of grating and plaid). Above: Fourier representation of the vertical grating (white dots) and the plaid components (black dots). Both fall within the same vertical detector (light ovals), but only the plaid components fall within the horizontal detector (dark ovals).

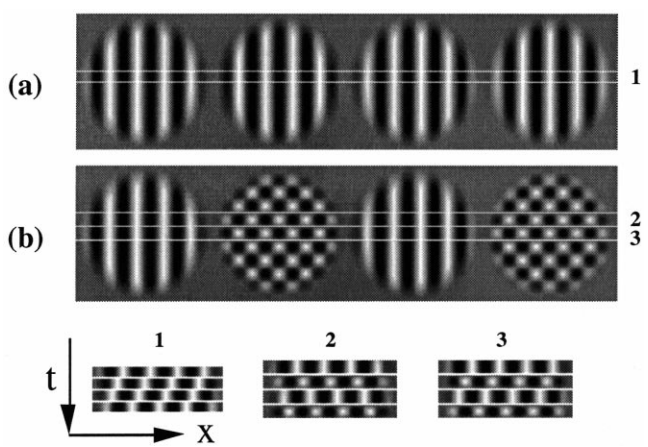


Fig. 6. Local phase shifts in the grating-plaid sequence. Experiment 2. Consecutive frames of two motion sequences; spatial phase steps by 90° from frame to frame. (a) Vertical grating; (b) vertical grating alternates with a plaid. Numbered strips are stacked to produce space–time diagrams in which time is increasing down the $y$-axis. Strip 1 drifts to the left, while strips 2 and 3 (adjacent in the grating-plaid sequence) drift to the right and left, respectively.

had a direction orthogonal to its neighbours. Along the diagonals of the checkerboard, adjacent squares have opposite directions which presumably cancel each other, or lose out in the competition with the coherent horizontal (or vertical) motion available in rows (or columns).

This is a reasonably satisfying account of an intriguing motion phenomenon, but for revealing the orientation sensitivity of the detectors the stimulus sequence used here is problematic. Different detectors are recruited at large and small angles, and so there is no way to test the responsiveness of the horizontal motion detector when $\theta > 45$. It was therefore necessary to devise a stimulus sequence that would anchor the motion to the horizontal direction and test only one detector.

## 3. Experiment 2: the grating-plaid sequence

To restrict motion to just one axis, we used the *grating-plaid sequence* (Fig. 5). The critical feature of this sequence is that the alternate frames of vertical grating prevent the mechanism for vertical motion (dark shaded ovals) from detecting anything other than flicker, as this mechanism is sensitive only to the counterphasing plaid components. The mechanism for horizontal motion (light shaded ovals), on the other hand, is able to detect both the plaid and grating components, and can therefore produce motion signals by integrating these two images over time. Initial observations confirmed that this sequence produced horizontal, but not vertical, strips of apparent motion (see Fig. 6).

From Fig. 5 it can be seen that increasing the angle ($\theta$) of the plaid components from the vertical while maintaining constant horizontal periodicity ($u$) has the effect of increasing the vertical spatial frequency ($v$). This means that, as before, the strips of motion are thinner with increasing $\theta$, but now the height of the strips is $1/(2\ v)$. At sufficiently high values of $\theta$, the plaid components will fall outside the detector (light grey ovals in Fig. 5) and motion perception will fail. The limiting value, $\theta_{max}$, gives a measure of the orientation range of that detector.

The height of the detector's RF may also be inferred from the grating-plaid sequence. We assume that when equal and opposite amounts of motion energy fall within the receptive field the response of that detector will be zero. It follows that when two oppositely moving (but otherwise identical) strips exactly cover the detector, no motion signal will arise (see Fig. 7b). This is not so if less than two strips fall on the detector (Fig. 7a), in which case there is a predominant motion signal in one direction, but it remains so with more than two strips (Fig. 7c). Thus at $\theta_{max}$, the height of the strips in the stimulus sequence should equal half the height of the motion detector. This strip height is the *critical strip height*.

### 3.1. Method

Apparatus was similar to experiment 1. The interleaved grating-plaid sequence had a temporal frequency of 7.5 Hz, and was viewed for 266 ms (two cycles). The

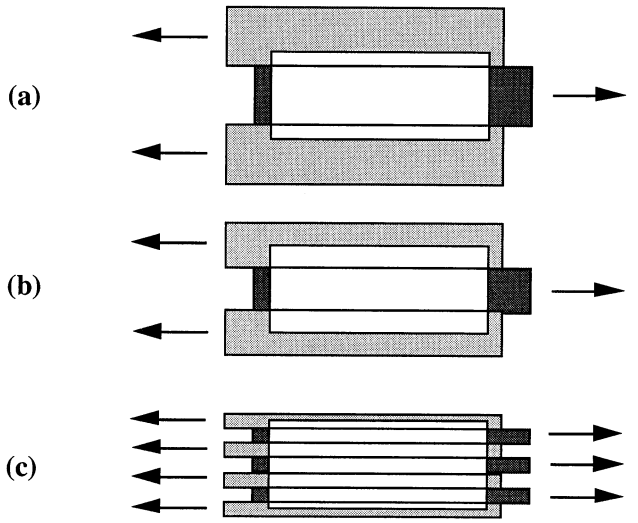


Fig. 7. Inferring the receptive field height. White rectangle represents the receptive field (RF) of a motion detector; light and dark grey bars are leftward and rightward strips of the grating-plaid sequence. (a) Most of the RF is occupied by a rightward-moving strip, and the detector signals rightward motion. (b) Equal areas of the RF are occupied by rightward- and leftward-moving strips, yielding an ambiguous signal. (c) Similar to (b) but more strips occupy the RF. Thus when strip height decreases to the *critical strip height* (b) motion signalling breaks down. RF height is given by twice the critical strip height (but see text for exact definition of height $H$).

stimulus spatial frequencies were changed by adjusting the viewing distance across sessions; the four viewing distances (20, 71, 214 and 642 cm) gave horizontal spatial frequencies of 0.10, 0.33, 1.00 and 3.00 cpd for both grating and plaid components. The horizontal periodicity of all stimuli was 4.5 cycles/image. Observers were required to indicate the direction of motion of a central strip in a single interval, binary choice task. This centrally located strip varied in height according to the vertical spatial frequency ($v$) of the plaid elements of the motion sequence, which varied across trials within a single session. A central fixation point was provided and cursors on each side of the display indicated the location and height of the central strip. Both the fixation point and side cursors were displayed for 250 ms before each trial. The contrast of the single gratings was fixed at 20% but the plaid contrast varied across sessions, taking values of 20, 40 or 80%. This was to compensate for any possible attenuation of the plaid components due to their off-vertical orientation and higher spatial frequency [$f=\sqrt{(u^2+v^2)}$]. Percent correct direction judgements were plotted against $\theta$ and limiting values $\theta_{max}$ were derived at 75% correct from a logistic curve fitted to the psychometric function data.

### 3.2. Results: motion receptive fields are very short

The results for observer NSS (see Fig. 8, left) show that the critical angle, $\theta_{max}$ at 75% correct, was greater than 80° at the lower spatial frequencies, decreasing to around 60° at the highest spatial frequency tested. This indicates that motion mechanisms are sensitive over a very broad range of orientations. Note also that when $\theta = 80°$ the radial spatial frequency ($f$) of the plaid's components is 5.8 times higher (2.5 octaves higher) than the grating, showing that the motion detector can combine information from components differing very widely in both orientation and spatial frequency. The key factor is the coherence of local phase changes. At all four spatial frequencies ($u$ values) there was no systematic effect of changing plaid contrast.

In general, $v/u = \tan\theta$, and so from the observed $\theta_{max}$ values we derive a corresponding $v_{max}$ which is the spatial frequency of the strips themselves (not the SF of the content of the strips) at the point where motion directions can just be resolved:

$$v_{max} = u \tan \theta_{max} \quad (1)$$

where $u$ and $v$ are the vertical and horizontal spatial frequencies of the plaid, and $\theta_{max}$ is the critical plaid angle (see Fig. 5). If we assume that the vertical profile of the receptive field is Gaussian then its extent can be represented as $2.5\sigma_y$, the Gaussian equivalent width, which is the width of a rectangle that has the same height and area as a Gaussian with standard deviation

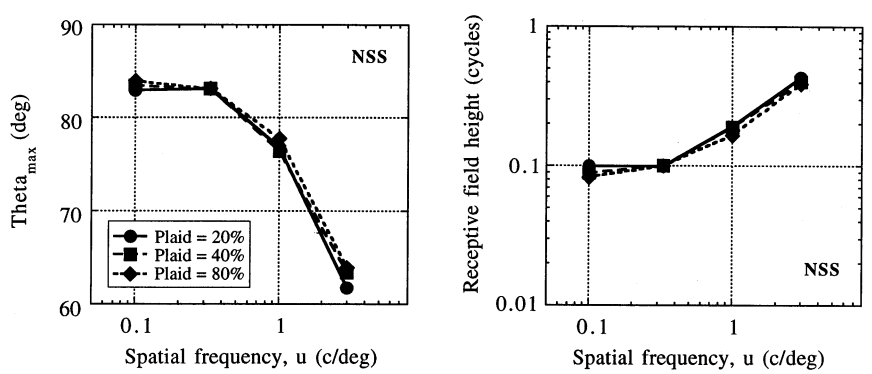


Fig. 8. Orientation range and RF height. Experiment 2. Orientation range in degrees (left) and RF height (H) in cycles (right) plotted against horizontal spatial frequency for one observer (NSS).

Table 1
Experiment 2: spatial and temporal frequency, speed and performance for NSS expressed in six different ways[a]

| | | | | |
|---|---|---|---|---|
| 0.1 | 0.33 | 1 | 3 | $u$, cpd |
| 7.5 | 7.5 | 7.5 | 7.5 | TF, Hz |
| 75 | 22.5 | 7.5 | 2.5 | Speed, deg/s |
| 84 | 84 | 77 | 63 | $\theta_{max}$, ° |
| 0.95 | 3.14 | 4.33 | 5.89 | $v_{max}$, cpd |
| 0.53 | 0.16 | 0.12 | 0.08 | Critical strip, deg |
| 0.42 | 0.13 | 0.09 | 0.07 | $\sigma_y$, deg |
| 0.08 | 0.08 | 0.18 | 0.41 | $H$, cycles |
| 50.45 | 15.29 | 11.08 | 8.15 | RF height, min arc |

[a] Bottom row shows RF height estimated as 2 $\sigma_y$ converted to min arc.

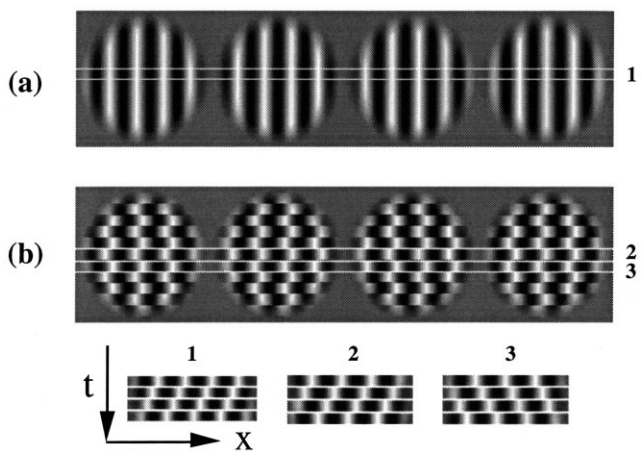


Fig. 9. Local phase in the strips stimuli. Experiment 3. Consecutive frames of two motion sequences. (a) Vertical grating; the four frames have phases 0, 90, 180 and 270°, reading from left to right; (b) is constructed from two oppositely-drifting gratings with phases 0, 90, 180 and 270°, and 270, 180, 90 and 0°, which are cut up horizontally and reassembled as alternate strips. The numbered slices are stacked to produce space–time diagrams in which time is increasing down the $y$-axis. Strip 1 drifts to the left, while 2 and 3, adjacent strips in sequence (b), drift to the left and right, respectively.

$\sigma_y$. The angle subtended by two strips is one vertical period of the plaid ($1/v$), and thus $\sigma_y$ can be estimated from the assumption outlined above that the equivalent height of the Gaussian equals two critical strip heights:

$$2.5\,\sigma_y = \frac{1}{v_{max}} \quad (2)$$

Combining Eqs. (1) and (2) we get:

$$\sigma_y = \frac{4}{10\,u\,\tan\theta_{max}} \quad (3)$$

Anderson and Burr (1991) expressed their RF height estimates ($H$) in cycles of the tested spatial frequency, as $H = 2\,u\,\sigma_y$, and so for direct comparison we have expressed our results the same way. Receptive field height ($H$, in cycles) is plotted against horizontal spatial frequency ($u$) in Fig. 8, right. $H$ was not constant, but increased with spatial frequency. A striking feature of the results, summarized in Table 1, was the shortness of the receptive fields, which ranged from as little as 0.08 cycles at 0.1 and 0.3 cpd, to 0.4 cycles (8 min arc) at 3.0 cpd.

### 3.3. Discussion: extraneous cue to motion, and the strip uncertainty problem

The second observer (MAG) reported the existence of an extraneous cue to direction in the grating-plaid sequence. An image-processing analysis of the grating-plaid sequence revealed that visual temporal integration over consecutive frames gave rise to an extraneous static cue to the direction of motion of the central strip. The cue had the appearance of a stationary, spatially periodic 'wiggly' distortion across the display in the vertical direction, with spatial frequency $v$. The spatial phase of this distortion signalled the direction of motion of the central strip. MAG's data were therefore discarded. The data from NSS, who did not notice this effect, were apparently unaffected by it.

The experiment was designed to reveal the orientation range and receptive field height of 1st-order motion detectors, but the decline in performance with decreasing strip height that was used to estimate these values could have been influenced by uncertainty as to which strip was to be judged. Uncertainty over which strip was being marked by the cursors could make performance fall to chance prematurely. It is therefore possible that the true receptive field height is even shorter (and the orientation range wider) than our estimates. Furthermore, if the filter were polar separable in the $u$–$v$ domain, rather than Cartesian separable, then this would also result in an underestimation of the receptive field height. The results must therefore be regarded as conservative estimates of these quantities.

## 4. Experiment 3: the moving strips sequence

To eliminate the artefactual cue described above, a third sequence was constructed by taking two drifting gratings, one moving to the left, the other to the right, and cutting both into horizontal strips of a given height. These strips were then stacked on top of one another, alternate strips coming from one or other of the two original gratings, to form the *strips sequence* (Fig. 9).

Comparison of Fig. 9 with Fig. 6 reveals that the local space–time diagrams of the grating-plaid and strips sequences are very similar. In fact, by shifting the relative phases of the two original gratings, it is possible to make a strips sequence that is practically indistinguishable from the grating-plaid sequence. In Fig. 9b, the phases used for the two constituent gratings are

0–90–180–270° and 270–180–90–0°. This arrangement ensures that adjacent strips are always 90° out of phase. An image-processing analysis of the new sequence confirmed that it did not contain a static cue like that found in the grating-plaid sequence.

### 4.1. Method

The apparatus and procedure were the same as experiment 2. The only difference was that the alternating frames of grating and plaid were replaced by consecutive frames of the strips sequence, and that all frames of the sequence were displayed at 20% contrast. The two authors and a naive volunteer acted as observers. Critical strip heights (at 75% correct) were estimated from the fitted psychometric functions, and converted into RF height estimates ($H$) as before.

### 4.2. Results: confirmation of the grating-plaid data

Fig. 10 shows the receptive field height estimates (H, in cycles) for all three observers, along with the data for observer NSS from experiment 2 (the grating-plaid sequence). Although there were slight variations across observers, the pattern of results was similar. The results from the strips sequence for NSS were encouragingly similar to the data from the grating-plaid sequence, confirming that the extraneous cue in experiment 2 had not affected NSS's data. As with the grating-plaid sequence, the most striking feature of the results is the shortness of the estimated receptive field heights, varying from around 0.1 to 0.4 cycles of the grating period over the spatial frequency range tested (0.1–3.0 cpd). The orientation range implied by these receptive field heights is extremely wide.

RF height estimates for observer NSS showed a slight but consistent decrease at each spatial frequency

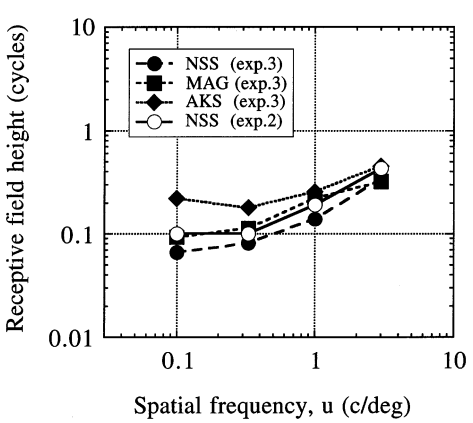


Fig. 10. Grating-plaid and strips sequences compared. Results for three observers (AKS, MAG and NSS) for experiment 3 (the strips sequence), and one observer (NSS) for experiment 2 (the grating-plaid sequence). Receptive field height $H$ in cycles is plotted against horizontal spatial frequency on log–log axes.

in experiment 3, compared with experiment 2. This may be related to the greater sharpness of the motion boundaries in the strips sequence, compared with the grating-plaid sequence.

## 5. Experiment 4: controlling for contrast and temporal frequency

In experiment 2, the grating frames had 20% contrast while the contrast of the plaid frames was 20, 40 or 80%; the level of plaid contrast had little or no effect upon direction discrimination performance. In experiment 3 the strips sequence was displayed at 20% contrast only. However, it is possible that performance might improve with increasing contrast because more energy is put into the motion mechanism. The critical strip height might decrease, and the orientation range would increase. Such contrast dependence would undermine our claim to be measuring RF height. Another potentially important variable is temporal frequency: the height estimates in experiments 2 and 3 were obtained at 7.5 Hz drift rate. It might be that larger RFs are recruited at higher speeds, in which case our estimates would increase with temporal frequency.

### 5.1. Method

The methods were as experiment 3, except that spatial frequency was 1.0 cpd in all cases, and across different sessions either (i) the contrast of the strips sequence was varied (5, 10, 20, 40 or 80%) at fixed temporal frequency (7.5 Hz), or (ii) the temporal frequency of the stimulus sequence was varied (15.0, 7.5, 3.75 or 1.875 Hz) at fixed contrast (20%). The former condition was the *contrast condition*, the latter the *temporal frequency condition*. The duration of the display was constant at 266ms in all cases, and in condition (ii) this meant that only half a temporal cycle was shown at 1.875 Hz. To prevent the motion sequence from appearing 'jerky' (undersampled), 45° phase shifts between consecutive frames were used at 1.875 Hz instead of the 90° phase shifts employed at the other temporal frequencies. Both authors acted as observers in the contrast condition (i), NSS in the temporal frequency condition (ii).

### 5.2. Results

(i) RF height estimates (H, in cycles) for the two observers (MAG and NSS) in the contrast condition are shown in Fig. 11(a). At contrasts below 20% H increased markedly, whereas at and above 20% contrast H was fairly constant. (ii) There was no systematic variation in critical strip height across temporal frequency (Fig. 11b), implying that the value of 7.5 Hz used in the earlier experiments was appropriate.

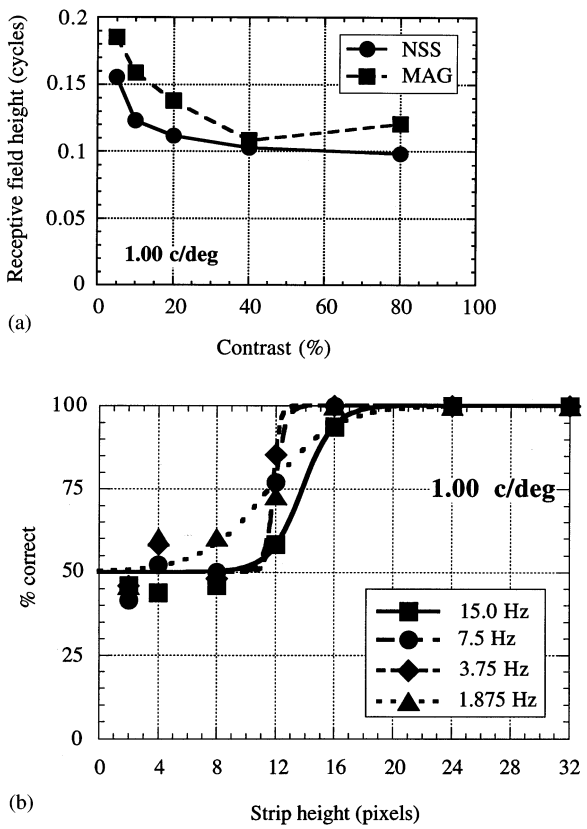


Fig. 11. The effect of contrast and temporal frequency. (A) Experiment 4(i). Receptive field height $H$ in cycles plotted against the contrast of the strips sequence for two observers (MAG and NSS) at a spatial frequency of 1.00 cpd. (B) Experiment 4(ii). Direction discrimination performance against strip height (in pixels) for NSS at four speeds (temporal frequencies) at 1 cpd, 20% contrast. 1 pixel = 0.53 min arc. Threshold strip height was about 6 min arc at all speeds.

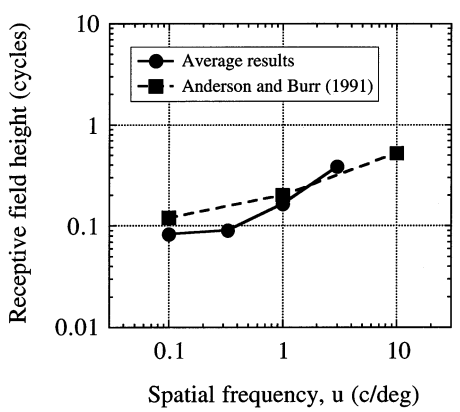


Fig. 12. RF height estimates. Averaged results for one subject (NSS) from experiments 2 and 3 (grating-plaid and strips sequences), compared with data from Anderson and Burr (1991). RF height in cycles is plotted against horizontal spatial frequency.

### 5.3. Discussion: previous parameter choices justified

The results from these two control experiments indicate that the values of RF height from experiments 2 and 3 (Fig. 10) are representative estimates. Since critical strip height reached its asymptotic level near 20% contrast, this finding validates the results from experiment 3 carried out at 20% contrast. RF height estimates would not have been much smaller at higher contrasts. This suggests that a structural limit, imposed by the receptive field height itself, has been reached at about 20–40% contrast. At lower contrasts reduction of signal:noise ratio may cause performance to be worse than this structural limit.

Variation of temporal frequency had no systematic effect on receptive field height estimates, at least up to 15Hz as tested here. Thus spatial resolution for shearing motion is determined by local spatial frequency ($u$), not by speed of movement. This conclusion contrasts with the findings of Koenderink et al. (1985) who found that the strip width required for spatial segregation of strips of random dots moving in opposite directions increased markedly with speed, being about 8 min arc at 1°/s but 100 min arc at 10°/s. The difference may be largely due to differences in the type of motion: shearing motion (along the strips) in our experiments versus expansive/compressive motion (across the strips) in the Koenderink study. As Koenderink et al. point out, the segregation limit in their experiment was probably temporal rather than spatial: the dots had to spend about 70–100 ms traversing one strip in order for the strips to be segregated. In an analogous study using shearing motion of drifting random dots, Nakayama and Silverman (1984); their Fig. 2) found that strips of motion only 13 min arc high could just be resolved at 16°/s (Nakayama & Silverman, 1984) which is much more in line with our findings.

## 6. General discussion

Fig. 12 shows a summary of the average RF height estimates from experiments 2 and 3 for observer NSS, along with those of Anderson and Burr (1991) who used a spatial summation technique to measure both the width and height of motion receptive fields at threshold. These authors published similar results from masking experiments (Anderson & Burr, 1987, 1989, 1991; Anderson, Burr & Morrone, 1991), and confirmed that the RF size (in cycles) increases with increasing spatial frequency. They also showed that the heights and widths of the receptive field envelopes were equal. There is excellent agreement between Anderson and Burr's (1991) results and our own, despite the very different methods and theoretical assumptions used. At all spatial frequencies the estimated receptive field height is a small fraction of one spatial period. The heights range from 0.08 cycles at 0.1 cpd to 0.38 cycles at 3.0 cpd, corresponding to orientation ranges (in the

plaid experiment) from about $\pm 84$ to $\pm 65°$, respectively.

### 6.1. The Gaussian derivative model

Do the estimates given above for RF height and orientation range correspond to plausible space–time filters in motion analysis? To begin to answer this question we consider a simple, general model for motion filtering. Direction-selective linear filters have receptive fields that are oriented in space–time (Adelson & Bergen, 1985) and it is well known that such fields can be formed by linear combination of pairs of RFs that are 90° out of phase with each other in space and time. This approach gives a good account of direction-selective simple cells (McLean & Palmer, 1989) and complex cells (Emerson, Bergen & Adelson, 1992) in the primary visual cortex. An analytically tractable form for this kind of model is based on Gaussian derivatives (GDs) in space and time (Young, 1985; Adelson & Bergen, 1986; Johnston, McOwan & Buxton, 1992; Young & Lesperance, 1993). Thus if $G(x, y, t)$ is a Gaussian function of space and time then $G_x$ and $G_t$ (where subscripts denote partial differentiation with respect to a given variable) are suitable subunits that form an oriented space–time receptive field ($M_1$) when added together: $M_1 = G_x + G_t$. Differentiating again with respect to $x$ gives us $M_2 = G_{xx} + G_{xt}$, which forms a direction-selective filter whose SF bandwidth is narrower than that of $M_1$. See Bruce, Green and Georgeson (1996) for illustration (their Fig. 8.7). Clearly, differentiating $m$ times with respect to $x$ can define a family of filters $M_m$ ($m = 1,2,3,4, \ldots$). $M_1$ and $M_2$ are odd and even filters in space–time and can be used as elements of the Adelson–Bergen energy model to encode velocity exactly. There is an equivalence between this version of the energy model and at least one form of the multichannel gradient model (Adelson & Bergen, 1986; Johnston et al., 1992; Bruce et al., 1996).

We now derive the spatio-temporal frequency response of $M_m$ in order to see how such filters would respond to the grating-plaid sequence of experiment 2. Since motion filters are not space–time separable, the 2-D spatial frequency response depends on temporal frequency, and the equations must be developed in 3-D (space–time) not 2-D space alone. Thus we define:

$$M_m(x, y, t) = \frac{\partial^m G}{\partial x^m} + \frac{\partial^m G}{\partial x^{m-1} \partial t} \quad (4)$$

where $G(x, y, t)$ is a Gaussian function with standard deviations $\sigma_x$, $\sigma_y$, $\sigma_t$. The Fourier transform $F(u, v, w)$ of $G(x, y, t)$ is:

$$F(u, v, w) = \exp[-2\pi^2(u^2 \sigma_x^2 + v^2 \sigma_y^2 + w^2 \sigma_t^2)] \quad (5)$$

Differentiation with respect to x or t multiplies the transform by $2\pi i u$ or $2\pi i w$, respectively. Hence the Fourier transform ($F_m$) of $M_m$ is:

$$F_m(u, v, w) = (2\pi i)^m u^{m-1}(u + w)F(u, v, w) \quad (6)$$

Fig. 13 shows iso-sensitivity contours of $F_2(u, v, w)$ in the $(u, v)$ plane for $w = 7.5$ Hz, plotted using the values of $\sigma_y$ from Table 1, and assuming $\sigma_x = \sigma_y$, $\sigma_t = 0.01$ s. These plots give an idea of the 2-D range of spatial frequencies that could be combined by the second-derivative based motion filter ($G_{xx} + G_{xt}$) in order to extract the motion of strips in the grating-plaid experiment. Positive $u$ values represent motion in the preferred direction, negative $u$ values are the non-preferred direction. The outermost contour in each plot is the locus of spatial frequencies $(u, v)$ for which the filter has 5% of its peak sensitivity at $w = 7.5$ Hz. We note that in each case the range of plaid frequencies ($v_{max}$) that supported perception of moving strips would be nearly accounted for if the limit of performance was imposed by the 5% sensitivity contour.

For example, our empirical estimate of $\sigma_y$ was 0.09° for a grating of 1 cpd; Fig. 13(C) shows that the $M_2$ filter defined by this spatial scale has a peak sensitivity at $u = 1.9$ cpd, and at $u = 1$ cpd the 5% contour extends out to $v = 4.0$ cpd. This compares well with our experimental $v_{max}$ of 4.33 cpd (Table 1). Across the four test spatial frequencies (0.1, 0.33, 1, 3 cpd) the $v_{max}$ values from Table 1 are 0.95, 3.1, 4.3, 5.9 cpd, compared with $v = 0.78$, 2.5, 4.0, 5.7 cpd at the 5% limit derived from Fig. 13(A–D). The agreement here is fairly good, but the argument depends on the assumption that human performance fails when the filter's response to one plaid component falls to about 5% of its response to the vertical grating (given that the SF of the vertical grating lies near the peak of the filter). Bearing in mind that there are two plaid components, the limiting response (to the plaid) should be about 10% of the response to the optimal grating. We have independent evidence that this limit is about right. In a separate study (Georgeson & Scott-Samuel, 1999) the motion sequence (cf. Fig. 1) consisted simply of interleaved 1 cpd vertical gratings of two different contrasts. Recognition of direction fell to chance when the contrast of one grating fell below about 10% of the other grating, irrespective of absolute contrast. This would correspond to a 5% response limit for each of the plaid's components.

In summary, then, the 2-D spatial bandwidth of the second-derivative motion filter ($M_2$) based upon a circular Gaussian (where $\sigma_x = \sigma_y$) appears well suited to account for our results on spatial resolution of strips of motion. The similarity between our RF height estimates and those of Anderson and Burr (1991) obtained by very different means lends strong support to this conclusion. Fig. 14 illustrates the short, stubby receptive fields of the $M_2$ filter that are implied by this analysis.

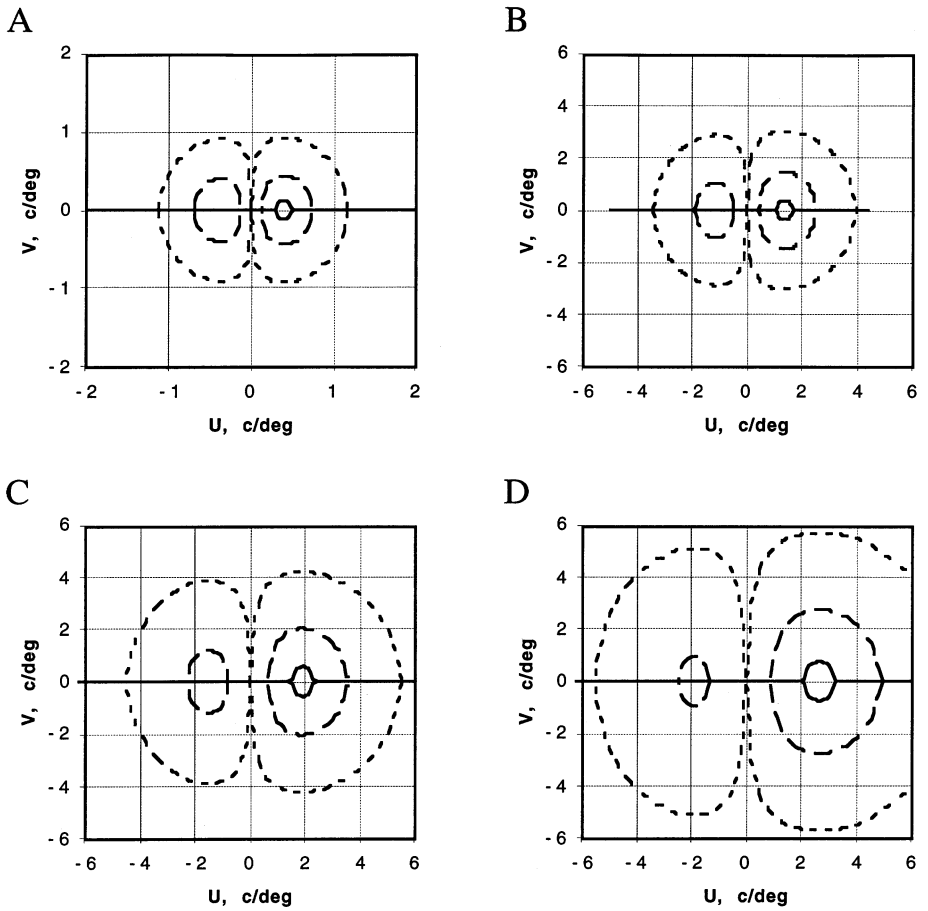


Fig. 13. Gaussian derivative filters for motion. Contour maps of the 2-D spatial frequency response in the $(u, v)$ plane of the space–time oriented filter $G_{xx} + G_{xt}$ (see text) at a temporal frequency $w = 7.5$ Hz. Solid curve: 95% of peak sensitivity; long dashes: 50%; short dashes: 5%. The 95% contour exists only in the right half-plane ($u > 0$) representing the preferred direction of motion. (A) 0.1 cpd test. Gaussian parameters: $\sigma_x = \sigma_y = 0.42°$, $\sigma_t = 0.01$ s. (B) 0.33 cpd test; $\sigma_x = \sigma_y = 0.13°$. (C) 1.0 cpd test; $\sigma_x = \sigma_y = 0.09°$. (D) 3.0 cpd test; $\sigma_x = \sigma_y = 0.07°$. Note: the surface mapped here does *not* represent the Fourier transform of a spatial filter kernel, but is the sensitivity in a cross-sectional $(u, v)$ plane of the direction-selective spatio-temporal filter that is oriented in the $(u, w)$ plane.

### 6.2. Orientation tuning

The orientation tuning of the $M_m$ family of GD filters is easily derived from Eq. (6) by substituting $u = r \cos\theta$ and $v = r \sin\theta$, where $r = \sqrt{(u^2 + v^2)}$ is radial frequency and $\theta$ is orientation. Assuming $\sigma_x = \sigma_y$, this yields:

$$F_m(\theta, r, w) = K(r \cos\theta)^{m-1}(r \cos\theta + w) \quad (7)$$

where $K = (2\pi i)^m \exp[-2\pi^2(r^2\sigma_x^2 + w^2\sigma_t^2)]$ does not vary with orientation. Thus the shape of the orientation tuning curve (Eq. (7)) depends on $m$, $r$ and $w$, but not on spatial scale provided that $\sigma_x = \sigma_y$. Orientation tuning curves are usually plotted as a function of $\theta$ at a fixed value of spatial and temporal frequency $(r, w)$, and we see from Eq. (7) that for a static or slowly moving grating the function (normalized to a peak of 1.0) tends to $(\cos\theta)^m$ as $w$ tends to 0, but at high temporal frequencies if $w >> r$ then it tends to $(\cos\theta)^{m-1}$. The corresponding bandwidths $\theta_B$ (half-width at half-height) increase from $60^{1/m}$ to $60^{1/m-1}$°, respectively. Thus orientation tuning of the GD model filter broadens with increasing temporal frequency, as

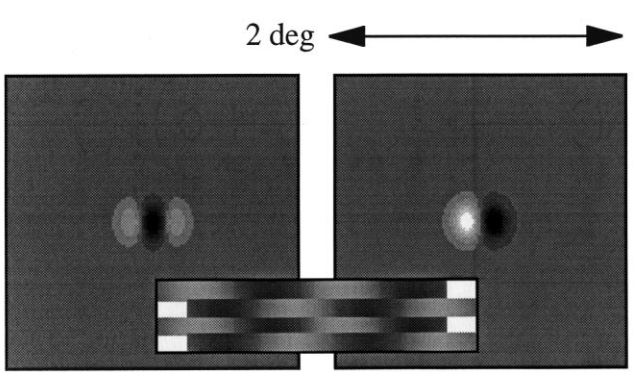


Fig. 14. Subunits of the motion filter at 1 cpd. Spatial receptive fields of the even ($G_{xx}$, left) and odd ($G_x$, right) subunits of the $M_2$ motion filter. Gaussian parameters were $\sigma_x = \sigma_y = 0.09°$, as used in Fig. 13(C). Inset, drawn to scale, are strips of a 1 cpd grating at the critical strip height observed experimentally.

observed in adaptation experiments (Snowden, 1992) and masking experiments (Anderson et al., 1991). The broadening of model tuning is greatest for $m = 1$ where the filter has a half-bandwidth $\theta_B$ of 60° when $w = 0$, but becomes completely non-oriented when temporal frequency is high. This correctly reflects the behaviour of adaptation and masking at very low spatial frequencies (0.1–0.2 cpd) (Kelly & Burbeck, 1987; Snowden, 1992). When $m = 2$, $\theta_B$ increases from 45 to 60° with increasing temporal frequency; when $m = 4$ it shifts from 32.8 to 37.5°, and when $m = 6$ it shifts from 27.0 to 29.5°. Thus increasing the derivative number ($m$) mirrors the two effects of adaptation observed by Snowden: with increasing spatial frequency the orientation bandwidths decrease, and the effect of temporal frequency on bandwidth progressively disappears. Experimental estimates of spatial frequency bandwidth (in octaves) also become narrower with increasing peak SF (Wilson, Levi, Maffei, Rovamo & De Valois, 1990) and so an attractive way to model this important cluster of results in the Gaussian derivative (GD) framework (Eq. (4)) is to suppose that for low SF channels $m = 1$ or 2, but at higher SFs $m$ increases progressively to around 4–6 (see Bruce et al., 1996, their Fig. 5.8). Accompanying the increase in $m$ is a decrease in the spatial scale ($\sigma_x$, $\sigma_y$).

Many details remain to be worked out, partly because the available data on bandwidths are fragmentary, and it also appears that Snowden's (1992) orientation bandwidths are systematically smaller than one would expect from the GD model with a circular base Gaussian (i.e. $\sigma_x = \sigma_y$). Physiological data indicate that cells in visual cortex have quite a range of bandwidths at each centre frequency (De Valois, Albrecht & Thorell, 1982), and so it is possible that different psychophysical tasks tap into different subsets of cells. For example, the tuning revealed in adaptation experiments might reflect the more narrowly tuned cells at each spatial frequency, while our experiments may reveal the shortest RFs and hence more broadly tuned cells. Although Fig. 13 illustrates the SF tuning surfaces for $m = 2$, our results are consistent with the idea outlined above that the derivative order $m$ increases with the peak SF of the filter. Setting $m = 3$ gave good predicted values of $v_{max}$ at 3 cpd, but not at 0.1 or 0.33 cpd where the narrowing of bandwidth (in going from $m = 2$ to $m = 3$) made the filters too insensitive at low SFs.

### 6.3. Off-frequency looking?

Our experiments yield fairly direct estimates of the RF length factor ($\sigma_y$), but not the width factor ($\sigma_x$). In analyzing the GD model we have supposed that $\sigma_x = \sigma_y$, and shown that this model is consistent with our data. What happens if $\sigma_x << \sigma_y$? This creates a filter with a more conventional, elongated but narrower RF, and a higher peak SF with narrower orientation tuning at the peak SF. However, it turns out that in the 2-D frequency domain (as Fig. 13) the 'skirts' of the $M_2$ filter response can be quite broad, and that an oriented filter centred at say 4 or 6 cpd could have sufficiently broad orientation bandwidth at $v = 1$ cpd to be consistent with our 1 cpd data, though the sensitivity to the grating component would be low. Such use of an effective, but less sensitive channel is known as 'off-frequency looking'. It could play a part in our experiments, where contrast is well above threshold. However, off-frequency looking is unlikely to contribute much to performance at contrast threshold, and so the similarity between our results and those of Anderson and Burr (1991) obtained at threshold (Fig. 12) strengthens the argument that short, stubby receptive fields for motion do exist in human vision (Fig. 14) (though for a different view, see Watson and Turano (1995)).

### 6.4. Implications for motion integration in plaids

The very broad orientation selectivity implied by these short RF heights suggests a surprisingly simple solution to the aperture problem for coherent plaid motion. Such broadband detectors would passively integrate plaid components over a wide range of orientations, and hence respond to the movement of local structure (local phase shifts) in the plaid, as we have shown. Active schemes for combining the motion components of a plaid pre-suppose that the components are initially processed separately in narrowband filters. Our results suggest that this may not be true in general, and so more explicit schemes for combination such as the intersection of constraints (IOC) rule (Adelson & Movshon, 1982) or vector–sum rule (Wilson, Ferrera & Yo, 1992) might be unnecessary in many cases. Perhaps we should not be asking how component motions are combined, but how are they ever segregated? Recent models (Wilson & Kim, 1994) and experiments (Smith, Curran & Braddick, 1999) have begun to address this issue.

## Acknowledgements

This work was supported by a BBSRC studentship to NSS, and by BBSRC grant GR/G63582 to MAG. Parts of this work have previously been presented in abstract form (Georgeson & Scott-Samuel, 1994; Scott-Samuel & Georgeson, 1996).

## Appendix A

For experiment 1, the height of the strips seen in the stimulus sequence can be derived as follows. A sinusoidal grating may be written as:

$$G(x, y) = \sin[2\pi(u x + v y) - \phi] \quad \text{(A1)}$$

where $u$ and $v$ are its horizontal and vertical spatial frequencies (which together determine its orientation, $\theta = \arctan(v/u)$), and $\phi$ is its phase. The local phase, $\phi'$, expressed relative to an arbitrary position $(x_1, y_1)$ is given by:

$$\phi'(x_1, y_1) = \phi - 2\pi(u x_1 + v y_1) \quad \text{(A2)}$$

So, for a 0–90–180–270° sequence with gratings tilting alternately one way and then the other (where the horizontal spatial frequency $u$ is constant and the vertical spatial frequency alternates between $+v$ and $-v$) the local phases in each of the four frames are given by:

$$\phi'_1 = -2\pi(u x_1 + v y_1) \quad \text{(A3)}$$

$$\phi'_2 = \frac{\pi}{2} - 2\pi(u x_1 - v y_1) \quad \text{(A4)}$$

$$\phi'_3 = \pi - 2\pi(u x_1 + v y_1) \quad \text{(A5)}$$

$$\phi'_4 = \frac{3\pi}{2} - 2\pi(u x_1 - v y_1) \quad \text{(A6)}$$

We propose that the local direction of perceived motion depends on the direction of local phase shift. Thus motion to the right occurs when:

$$\phi'_1 < \phi'_2 < \phi'_3 < \phi'_4 \quad \text{(A7)}$$

and the reverse is true for motion to the left. To determine the height and direction of the strips perceived in the sequence, it is necessary to find the range of $y$ values over which there is a common direction of motion, i.e. within which $\phi'_1 < \phi'_2$ and $\phi'_2 < \phi'_3$. From Eqs. (A3), (A4) and (A5) we require:

$$-2\pi(u x_1 + v y_1) < \frac{\pi}{2} - 2\pi(u x_1 - v y_1) \quad \text{and}$$

$$\pi - 2\pi(u x_1 + v y_1) > \frac{\pi}{2} - 2\pi(u x_1 - v y_1) \quad \text{(A8)}$$

which simplifies to:

$$y_1 > \frac{-1}{8v} \quad \text{and} \quad y_1 < \frac{1}{8v} \quad \text{(A9)}$$

Thus the range of $y$ values that share a common direction of phase shift is $\pm 1/(8v)$, and the height of such moving strips is $1/(4v)$, one quarter of a period of $v$.

## 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.

Adelson, E. H., & Bergen, J. R. (1986). The extraction of spatio-temporal energy in human and machine vision. IEEE workshop on motion: representation and analysis, Charleston, S.C.

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*, 523–525.

Anderson, S. J., & Burr, D. C. (1987). Receptive field size of human motion detection units. *Vision Research*, *27*, 621–635.

Anderson, S. J., & Burr, D. C. (1989). Receptive field properties of human motion detector units inferred from spatial frequency masking. *Vision Research*, *29*, 1343–1358.

Anderson, S. J., & Burr, D. C. (1991). Spatial summation properties of directionally selective mechanisms in human vision. *Journal Of the Optical Society Of America A*, *8*, 1330–1339.

Anderson, S. J., Burr, D. C., & Morrone, M. C. (1991). Two-dimensional spatial and spatial-frequency selectivity of motion-sensitive mechanisms in human vision. *Journal of the Optical Society of America A*, *8*, 1340–1351.

Braddick, O. J. (1993). Segmentation versus integration in visual motion processing. *Trends in Neurosciences*, *16*, 263–268.

Bruce, V., Green, P. R., & Georgeson, M. A. (1996). *Visual perception: physiology, psychology and ecology* (3rd ed.). Hove, UK: Psychology Press.

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

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*, 545–559.

Emerson, R. C., Bergen, J. R., & Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, *32*, 203–218.

Fredericksen, R. E., Verstraten, F. A. J., & van de Grind, W. A. (1997). Pitfalls in estimating motion detector receptive field geometry. *Vision Research*, *37*, 99–119.

Georgeson, M. A., & Scott-Samuel, N. E. (1994). Orientation bandwidth of motion perception. *Ophthalmic and Physiological Optics*, *14*, 438–439.

Georgeson, M. A. & Scott-Samuel, N. E., (1999). Motion contrast: a new metric for direction discrimination. *Vision Research*, *39*, 4393–4402.

Georgeson, M. A., & Shackleton, T. M. (1989). Monocular motion sensing, binocular motion perception. *Vision Research*, *29*, 1511–1523.

Gorea, A. (1985). Spatial integration characteristics in motion detection and direction identification. *Spatial Vision*, *1*, 85–102.

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 B*, *250*, 297–306.

Kelly, D. H., & Burbeck, C. A. (1987). Further evidence for a broadband, isotropic mechanism sensitive to high-velocity stimuli. *Vision Research*, *27*, 1527–1537.

Koenderink, J. J., van Doorn, A. J., & van de Grind, W. A. (1985). Spatial and temporal parameters of motion detection in the peripheral visual field. *Journal of the Optical Society of America A*, *2*, 252–259.

Ledgeway, T., & Smith, A. T. (1994). Evidence for separate motion-detecting mechanisms for first- and second-order motion in human vision. *Vision Research*, *34*, 2727–2740.

McLean, J., & Palmer, L. A. (1989). Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in area 17 of the cat. *Vision Research*, *29*, 675–679.

Nakayama, K., & Silverman, G. H. (1984). Temporal and spatial characteristics of the upper displacment limit for motion in random dots. *Vision Research*, *24*, 293–299.

Scott-Samuel, N. E., & Georgeson, M. A. (1996). Receptive-field length of motion sensors in human vision. *Perception* (*suppl.*), *25*, 125.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research*, *39*, 2853–2865.

Shadlen, M., & Carney, T. (1986). Mechanisms of motion perception revealed by a new cyclopean illusion. *Science*, *232*, 95–97.

Smith, A. T., Curran, W., & Braddick, O. J. (1999). What motion distributions yield global transparency and spatial segmentation? *Vision Research*, *39*, 1121–1132.

Snowden, R. J. (1992). Orientation bandwidth: the effect of spatial and temporal frequency. *Vision Research*, *32*, 1965–1974.

van den Berg, A. V., van de Grind, W. A., & van Doorn, A. J. (1990). Motion detection in the presence of local orientation changes. *Journal of the Optical Society of America A*, *7*, 933–939.

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

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

Wilson, H. R., & Kim, J. (1994). A model for motion coherence and transparency. *Visual Neuroscience*, *11*, 1205–1220.

Wilson, H. R., Levi, D. M., Maffei, L., Rovamo, J., & De Valois, R. L. (1990). The perception of form: retina to striate cortex. In L. Spillman, & J. S. Werner, *Visual perception: the neurophysiological foundations* (pp. 231–272). London: Academic Press.

Young, R. A. (1985). The Gaussian derivative theory of spatial vision: analysis of cortical cell receptive field line-weighting profiles. Warren, MI: General Motors Research Labs. GMR-4920.

Young, R. A., & Lesperance, R. M. (1993). A physiological model of motion analysis for machine vision. Warren, MI: General Motors Research Labs. GMR-7878.