PERGAMON

Vision Research 39 (1999) 113–130

Vision Research

A computational model of selective deficits in first and second-order motion processing

Colin W.G. Clifford, Lucia M. Vaina *

Brain and Vision Research Laboratory, Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA

Received 1 September 1997; received in revised form 21 December 1997

Abstract

Recent neurological studies of selective impairments in first and second-order motion processing are of considerable relevance in elucidating the mechanisms of motion perception in normal human observers. We examine the stimuli which have been used to assess first and second-order motion processing capabilities in clinical subjects, and discuss the nature of the computations necessary to extract their motion. We find that a simple computational model of first and second-order motion processing is able to account for the data. The model consists of a first-order channel computing motion at coarse and fine scales, and a coarse scale second-order channel. The second-order channel is sensitive to motion information defined by variations in luminance, contrast, spatial frequency and flicker. When elements of the model are disabled, its performance on either first or second-order motion can be selectively impaired in line with the neurological data.

Keywords: Computational model; Motion; Psychophysics; Second-order motion; Visual cortex

1. Introduction

Our visual systems are able to extract motion information from a variety of cues. First-order motion is defined by spatiotemporal luminance variations in the retinal image; second-order motion is defined by variations in stimulus properties such as contrast, flicker and spatial frequency (Chubb & Sperling, 1988). The fact that normal human subjects are able to perceive second-order motion is of considerable theoretical interest, as standard energy (Adelson & Bergen, 1985; van Santen & Sperling, 1985) and correlational (Reichardt, 1961) models of motion detection are sensitive only to first-order motion. Motion energy models can be thought of as detecting simple patterns of spectral power in the Fourier domain representation of an image corresponding to first-order motion. For this reason, the term 'Fourier motion' has often been used to refer to first-order motion, with 'non-Fourier motion' referring to second-order motion. Here we shall use the terms first and second-order motion Fig. 1.

While first-order motion mechanisms are blind to second-order motion, physiological studies provide clear evidence of neurons sensitive to second-order motion. Directionally selective neurons in primate area MT (Albright, 1992; O'Keefe, Carandini, Beusmans & Movshon, 1993; O'Keefe & Movshon, 1997), MSTd (Geesaman & Anderson, 1996) and STP (O'Keefe & Movshon, 1996), and cat areas 17 and 18 (Zhou & Baker, 1993), have been found to respond to motion defined by both luminance and contrast cues. Recent brain imaging studies point to areas outside of human MT/V5 as being selectively responsive to motion boundaries and second-order motion. The kinetic occipital region (KO) found by Dupont, De Bruyn and Vandenberghe (1997) to respond to kinetic contours has been suggested to correspond to V2v or V3A (Smith, Greenlee, Singh, Kraemer & Hennig, 1997). Human V3A has been reported to respond selectively to motion boundaries (Reppas, Niyogi, Dale, Sereno & Tootell, 1997), and in macaque monkeys V3 exhibits complex motion properties typically observed at the later stages of processing (Gegenfurtner, Kiper & Levitt, 1997).

How is second-order motion analysed? One possibility is that the detection of second-order motion is quite separate from and quite different to the detection of first-order motion. For instance, second-order motion might be mediated by a feature-tracking mechanism reminiscent of the long-range system proposed by Braddick (1974), rather than involving low-level motion

* Corresponding author. Tel.: +1 617 3532455; fax: +1 617 3536766; e-mail: vaina@bu.edu

0042-6989/98/$ - see front matter
PII: S0042-6989(98)00082-0

detectors. Another possibility is that second-order motion is detected by a mechanism or mechanisms similar to the one which extracts first-order motion. This second-order mechanism could either operate in parallel with the first-order mechanism (Chubb & Sperling, 1989; Wilson, Ferrera & Yo, 1992) or in series operating on the output of the first-order motion calculation (Zanker, 1993, 1996). If the organisation is parallel, then some form of initial pre-processing must render the motion signals in the second-order pathway visible to first-order mechanisms. Chubb and Sperling (1991) propose that motion extraction in the second-order pathway is preceded by a texture grabbing operation. This might take the form of low-pass (Werkhoven, Sperling and Chubb, 1993) or band-pass (Wilson et al., 1992) filtering, followed by rectification. Alternatively, it is been suggested that first and second-order motion are analysed by a single mechanism, either through the temporal coherence of first-order motion signals (Grzywacz, Watamaniuk & McKee, 1995) through the combination of filters computing derivatives of spatiotemporal image intensity (Johnston, McOwan & Buxton, 1992; Johnston & Clifford, 1995a), or through pre-processing by an asymmetric compressive nonlinearity (Taub, Victor & Conte, 1997).

Numerous psychophysical studies bear upon the issue of whether first and second-order motion are analysed by a single mechanism or by distinct mechanisms and if the mechanisms are distinct, whether they share common features. Table 1 contains a list of studies concerned with whether there are one, two, or many channels involved in the analysis of first and second-order motion. The list is intended to be representative, rather than exhaustive, and does not include papers on chromatic motion or plaids.

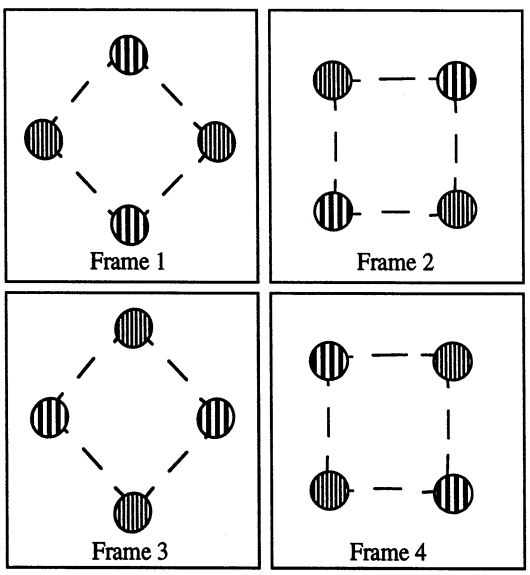

Fig. 1. Schematic diagram of the Green (1986) stimulus for generating second-order motion using Gabor patches. To perceive the motion of the Gabor patches the motion mechanism involved must be able to distinguish between patches of different spatial frequency.

While there is no clear consensus on the number of channels involved, the weight of psychophysical evidence appears to favour the existence of distinct first and second-order channels. Some of the most compelling evidence for the existence of specialised second-order motion mechanisms comes from the study of neurological patients. Plant and Nakayama (1993) report selective impairment to second-order motion in the visual hemifield contralateral to unilateral posterior cerebral lesions in three patients. In a study of 21 patients with unilateral brain damage, Greenlee and Smith (1997) found three patients with superior temporal or lateral inferoparietal lesions who had much higher threshold elevations for first-order speed discriminations than for second-order, but none with severely impaired second-order speed perception and intact first-order. Vaina and Cowey (1996) report the case of a patient, FD, who has a shallow lesion in the caudal portion of the superior temporal sulcus. FD shows a marked deficit in the perception of second-order motion in the hemifield contralateral to his lesion, but no impairment in his performance on psychophysical tasks with corresponding first-order stimuli. Vaina, Makris, Kennedy and Cowey (1998) also report the case of a patient, RA, who has a lesion superior to his calcarine fissure. RA shows a deficit in first-order motion, but his perception of second-order motion is unimpaired. Taken together, the visual deficits of patients FD and RA constitute a double dissociation of function between first and second-order motion processing. Vaina and co-workers interpret these patients' perceptual deficits as being due to disruptions of distinct first-order (RA) second-order (FD) motion processing channels.

In this paper we examine the psychophysical stimuli used in the above studies of first and second-order motion perception in neurological patients. We put forward a computational model of normal human vision able to extract motion in these stimuli. The model is a variant of those proposed by Chubb and Sperling (1989) and Wilson et al. (1992), consisting of distinct first and second-order channels. In the second-order channel the stimulus is pre-processed by a low-pass texture grabber (Chubb & Sperling, 1991; Werkhoven et al., 1993). Motion in both channels is extracted by a motion energy computation (Adelson & Bergen, 1985). When elements of the model are weakened or disabled, its performance on first or second-order motion can be selectively impaired to account for the patient data.

2. Neurological evidence for separate first and second-order motion processing

Normal human observers are able to perceive second-order motion defined by any of a range of stimulus

Table 1
List of psychophysical experiments, with normal or neurological subjects, bearing on the issue of how many channels mediate human perception of first and second-order motion

Authors	Task	Stimulus	Number of channels
Neurological subjects			
Greenlee and Smith (1997)	Detection/DoM	CM noise (static)	1 or 2
Plant and Nakayama (1993)	DoM/speed discrim.	Beats	2
Vaina and Cowey (1996)	DoM	Contrast/flicker/spatial frequency	2
Vaina, Makris, Kennedy and Cowey (1998)	DoM	Contrast/flicker/spatial frequency	2
Normal subjects			
Benton, Johnston and McOwan (1997)	Reverse Phi DoM	CM noise (static/dynamic)	1
Boulton and Baker (1991)	D_{max}/DoM	Gabor micropatterns	2
Chubb and Sperling (1988)	DoM	Contrast/flicker	2
Chubb and Sperling (1991)	DoM	Texture quilts	2
Cropper (1994)	Speed discrimination	Beats	2
Derrington and Badcock (1985)	Adaptation	Beats	2
Derrington, Badcock and Henning (1993)	DoM	Beats	2
Dosher, Landy and Sperling (1989)	Kinetic depth effect	CM or flickering dots	2
Edwards and Badcock (1995)	Motion coherence	Global random-dot motion	2
Gegenfurtner and Hawken (1996)	Speed matching	CM noise (static)	2
Gorea (1995)	Reverse Phi DoM	LM gratings	2
Green (1986)	DoM	Gabor patches	2
Harris and Smith (1992)	Optokinetic nystagmus	CM noise (static)	2
Johnston and Benton (1997)	Speed discrimination	CM noise (static/dynamic)	1
Johnston and Clifford (1995a)	Speed matching	CM gratings	1
Landy, Dosher, Sperling and Perkins (1991)	Kinetic depth effect	CM or flickering dots	2
Ledgeway (1994)	Adaptation	CM noise (static)	1 or 2
Ledgeway and Smith (1994)	DoM interleaved	CM and LM gratings	2
Ledgeway and Smith (1994)	Adaptation	CM Noise (Static)	1 or 2
Ledgeway and Smith (1995)	Speed Matching	CM noise (static)	1 or 2
Ledgeway and Smith (1997)	Adaptation/speed	CM noise (static)	1 or 2
Lu and Sperling (1995a)	DoM	Alternating features	3
Lu and Sperling (1995b)	Various	Modulations of static noise	3
Mather (1991)	Adaptation	Contrast-reversing bar	2
Mather and West (1993)	DoM	Interleaved RDKs	2
Mather and Tunley (1995)	DoM	Interleaved RDKs	2
McCarthy (1993)	Adaptation	CM gratings	2
Nam and Chubb (1997)	DoM	CM and LM noise (static)	2
Nishida (1993)	DoM	RDKs	2
Nishida, Ashida and Sato (1994)	Adaptation	Flicker/granularity	2
Nishida, Edwards and Sato (1997)	Motion contrast	CM noise (static/dynamic)	2
Nishida and Sato (1992)	Adaptation and DoM	Band-pass RDKs	2
Nishida and Sato (1995)	Adaptation	Beats/flicker/granularity	2
Pantle (1992)	DoM	CM grating	2
Papathomas, Gorea and Chubb (1996)	DoM	Interleaved luminance and texture	1 or 2
Petersik (1995)	DoM	Various apparent motion	2+
Smith (1994)	DoM	CM noise (static)	3
Smith, Hess and Baker (1994)	DoM	CM noise (static)	1 or 2
Smith and Ledgeway (1997)	Detection / DoM	CM noise (static/dynamic)	2
Solomon and Sperling (1994)	DoM	Full and half-wave dot patterns	3
Solomon and Sperling (1995)	DoM	CM noise (static)	1 or 2
Stoner and Albright (1992)	Coherence/transparency	Luminance-flicker plaids	1 or 2
Turano and Pantle (1989)	Adaptation	CM gratings	1
Victor and Conte (1990)	DoM	Various	2
Victor and Conte (1992)	Coherence/transparency	Non-Fourier plaids	1
Werkhoven, Sperling and Chubb (1993)	DoM	Gabor patches	2
Zanker (1993)	DoM	Motion-defined random textures	2
Zanker (1996)	Motion coherence	Wavy motion	2

parameters. Here we review, from a computational perspective, the evidence of neurological patients with selective deficits in their perception of first or second-order motion.

2.1. Motion from spatial frequency

Braddick (1974) proposed that two processes underlie human motion perception: a short-range process ope-

rating over small displacements and durations mediated by low-level mechanisms, and a long-range process involving higher-level processing. Long-range motion perception has been shown to depend on the spatial frequency content of the stimulus (Green, 1986; Boulton & Baker, 1991, 1993; Werkhoven et al., 1993; Werkhoven, Sperling & Chubb, 1994). Green (1986) used Gabor patch stimuli to study long-range apparent motion. His stimulus consisted of two pairs of Gabor patch stimuli of different spatial frequencies which would appear to rotate if correspondence could be established between similar patches, but would otherwise appear ambiguous in motion. Green found that performance on this task depended on the difference in carrier spatial frequency between the pairs, consistent with the operation of a second-order long-range mechanism (Green, 1986; Werkhoven et al., 1993). In a series of experiments, Boulton and Baker (1991; 1993) used a two-frame apparent motion stimulus made up of a number of identical Gabor micropatterns to investigate the parameters controlling motion perception. They found that at high stimulus densities the perceived direction of motion of the stimulus was governed by the motion of the carrier, suggesting the operation of a short-range first-order process, while with sparse stimuli perceived motion was dictated by the envelope, consistent with the operation of a long-range second-order process.

FD and RA were first tested in a replication of Green's experiment (Vaina & Cowey, 1996; Vaina et al., 1998). FD was impaired in all but the highest spatial frequency condition in the visual hemifield contralateral to his lesion, while RA's performance was within the normal range over the full set of spatial frequencies studied. RA was then tested in a replication of Boulton and Baker's experiment. His results failed to show the marked qualitative difference between sparse and dense stimulus conditions observed in normal observers. Instead, his perception of motion appeared to be governed largely by the motion of the envelope in all cases.

To perceive the motion of the Gabor patches in Green's stimulus requires that the motion mechanism involved be able to distinguish between tokens of different spatial frequencies. One way to achieve this would be to employ a pre-processing stage prior to motion extraction which converted signals of different spatial frequencies into signals of different strengths. Such a scheme was proposed by Werkhoven et al. (1993) to account for psychophysical data in an extension of Green's experiment. Werkhoven's scheme is a form of 'texture grabber' (Chubb & Sperling, 1991) realised by spatial low-pass filtering followed by rectification of the image signal. In this way variations in spatial frequency in the raw image give rise to variations in intensity in the pre-processed image, and are thus rendered visible to first-order motion analysis. Similarly, a pre-processing transformation from spatial frequency to intensity also recovers the envelope in the Boulton and Baker stimulus.

2.2. Motion from flicker

Motion can be perceived on the basis of temporal frequency information. For example, the direction of motion of a bar is distinguishable from a static random-dot background only through the flickering of a certain proportion of the dots making up the bar (Chubb & Sperling, 1988; Albright, 1992).

Both FD and RA were tested on flickering bar stimuli moving at various speeds. FD required about three times as many flickering elements as normal controls in order to discriminate the direction of motion of the bar when presented in the visual hemifield contralateral to his lesion, while RA's performance was normal.

To perceive the motion of the flickering bar requires that the motion mechanism be able to distinguish between regions on the basis of temporal frequency content. One way to achieve this would be to apply a band-pass or high-pass temporal filter to the image sequence, followed by a rectification, in order to pick out the bar from the background. The flicker in the bar would give rise to a strong signal after temporal filtering, while the static background would produce no signal. An alternative would be to apply a low-pass temporal filter to the image. In this case the background would produce the stronger signal after filtering.

2.3. Motion from contrast

In order to compare the operation of first and second-order motion processing, it is desirable to use stimuli which are closely matched in terms of appearance and spectral content. Plant and Nakayama (1993) and Greenlee and Smith (1997) used luminance-defined and contrast-defined gratings to probe the first and second-order channels respectively. Vaina and co-workers used random-dot patterns (Vaina & Cowey, 1996; Vaina et al., 1998).

Plant and Nakayama (1993) reported the case of a patient who suffered an impairment of second-order motion perception while performing normally on an equivalent first-order task (Plant, Laxer, Barbaro & Schiffman, 1993). Plant and Nakayama's 'Case 1' had suffered a unilateral posterior cerebral lesion involving lateral occipital and parietal cortex. Psychophysical testing revealed impaired motion perception in the hemifield contralateral to the lesion. Weber fractions for speed and temporal frequency discrimination thresholds for a first-order stimulus (a sine grating) were found to be substantially elevated. Contrast thresholds for direction-of-motion (DOM) discrimination were unaffected for sine grating stimuli, but elevated for contrast modulated

gratings. Case 1's unimpaired DOM discrimination thresholds for first-order stimuli show that his initial motion detectors were unaffected by the lesion. Plant and Nakayama argued that a selective deficit in the perception of contrast modulated gratings with no corresponding deficit in the perception of luminance modulated stimuli was evidence of a second stage of motion processing required for DOM discrimination with contrast modulated gratings. In order to account for Case 1's elevated Weber fractions with first-order stimuli they suggested that the region of extra-striate visual cortex required to carry out second-order DOM discriminations also facilitates accurate speed discrimination in first-order stimuli. Greenlee and Smith (1997) also report considerable overlap in the cortical areas involved in first and second-order speed perception.

In Vaina and co-workers' first-order random-dot pattern, moving luminance-defined elements are additively superimposed upon a dynamic random noise background (Vaina & Cowey, 1996; Vaina et al., 1998). In the second-order pattern, the dots differ from the background in contrast but not in luminance. The Vaina random-dot patterns and the Greenlee and Smith (1997) stimuli are similar in the sense that the first-order patterns are composed of a moving luminance signal added to a noise background, while the second-order patterns are composed of a moving envelope multiplying the contrast of the noise. However, to minimise possible first-order artefacts, the Vaina stimulus used a dynamic noise carrier while Greenlee and Smith used high-pass spatial filtering of static noise.

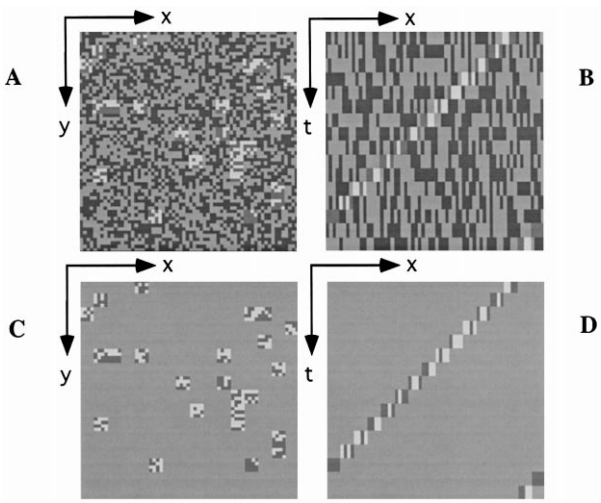

Fig. 2. (A) A single frame from a first-order random-dot pattern of the form used to test global motion perception in FD and RA (Vaina & Cowey, 1996; Vaina et al., 1998). The dots are distinct from the background only in their luminance. (B) A space-time slice through the image sequence showing the motion of a single dot against the dynamic noise background. (C) A frame from and (D) a space-time slice through the transformed image sequence after pointwise full-wave rectification showing that such an operation removes the noise from the background but not from dots themselves.

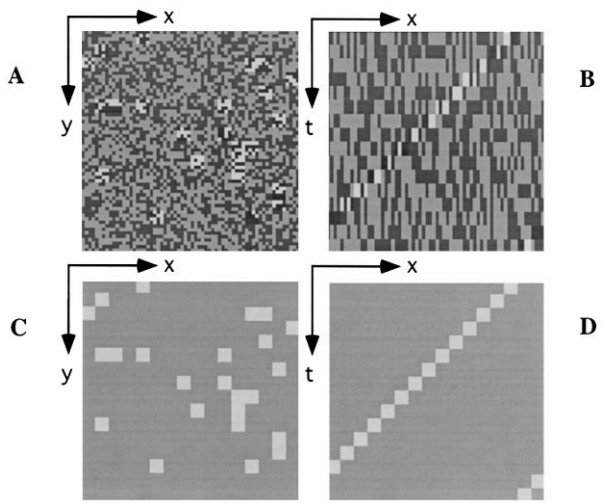

Fig. 3. (A) A single frame from a second-order random-dot pattern of the form used to test global motion perception in FD and RA (Vaina & Cowey, 1996; Vaina et al., 1998). The dots are distinct from the background only in their contrast. (B) A space-time slice through the image sequence showing the motion of a single dot against the dynamic noise background. (C) A frame from and (D) a space-time slice through the transformed image sequence after pointwise full-wave rectification showing that such an operation removes all the noise from stimulus, providing a purely first-order signal for subsequent motion analysis.

Fig. 2A shows a frame from an image sequence constructed in the same way as Vaina's first-order random-dot sequence, and Fig. 2B a space-time slice through the sequence. Fig. 3A, B show the corresponding views of a second-order sequence. The motion of the dots in the first-order stimulus is detectable by a first-order mechanism, although the additive dynamic noise tends to obscure the motion unless it is computed at coarse resolution. In the second-order stimulus, the noise acts as the carrier for the moving contrast-defined dots. There is on average no net luminance difference between the dots and the background in the second-order stimulus, so the motion of the dots is invisible to a first-order mechanism. Pointwise full-wave rectification of the second-order stimulus around the mean luminance removes all noise to reveal the underlying pattern (Fig. 3C) and motion (Fig. 3D) of the dots, rendering them visible to first-order motion detection. When the same operation is applied to the first-order stimulus, a component of dynamic noise remains which tends to mask the motion of the dots to subsequent first-order analysis (Fig. 2C, D).

Patients FD and RA were tested on a global motion task using the two random-dot stimuli. The task was a two alternative forced choice discrimination of the direction of motion of a population of dots. A variable proportion of the dots were moving in a common direction, while the remainder were randomly repositioned within the stimulus. The subjects' performance on the task was recorded as a function of the proportion of dots moving in a common direction, and a coherence

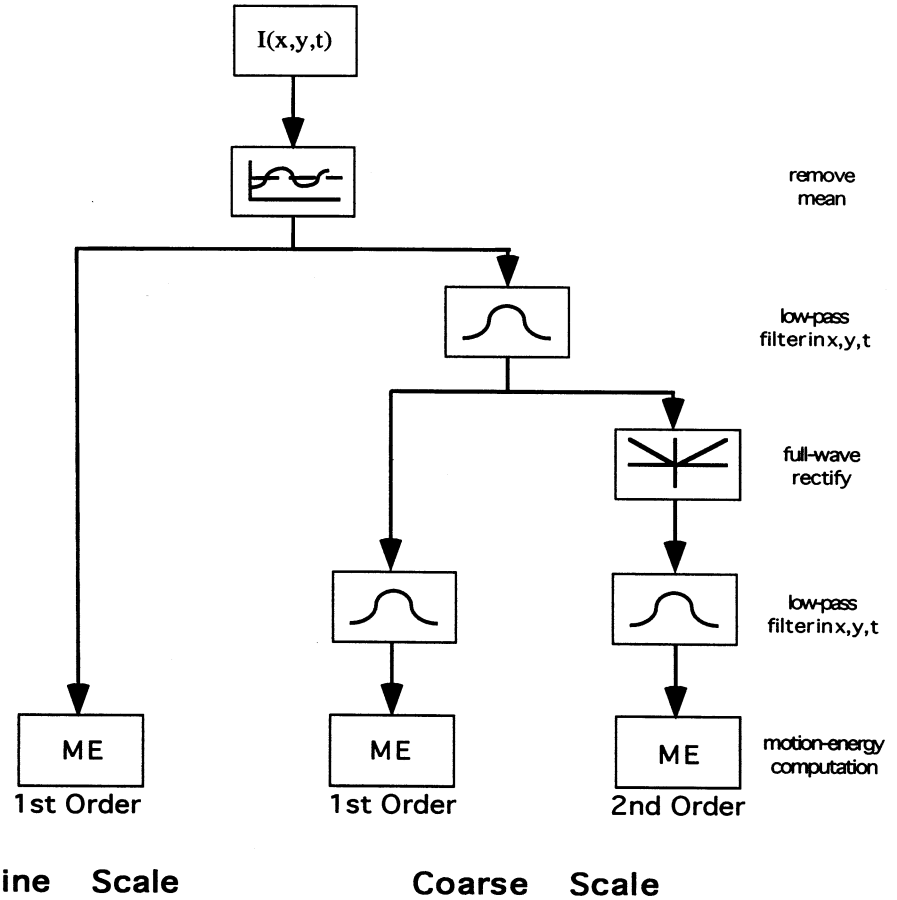

Fig. 4. Schematic diagram of the processing steps carried out in the model. The first-order channel analyses motion at both coarse and fine scales. The second-order channel operates only at the coarse scale. In all cases the first step is removal of the mean level of the stimulus. The fine scale part of the first-order channel operates on this mean zero version of the stimulus. In coarse scale motion analysis, the signal is then low-pass filtered in space and time. Prior to first-order motion analysis at the coarse scale the signal is further low-pass filtered. In the second-order channel, the signal is full-wave rectified and the mean again removed after the first stage of low-pass filtering. Processing then proceeds in the same way as for coarse scale first-order motion.

threshold calculated. Patient FD performed normally in both visual hemifields with first-order motion, but was badly impaired with second-order motion in the hemifield contralateral to his lesion. Patient RA performed normally with second-order motion, but was impaired with first-order motion in the hemifield contralateral to his lesion.

3. Model

We propose a two-channel model for the perception of first and second-order motion. The first-order channel processes motion information at coarse and fine spatial scales, while the second-order channel operates only at a coarse resolution. The first-order channel analyses motion according to a simple first-order luminance-based scheme. In the second-order channel the stimulus is first passed through a spatially and temporally low-pass filter and full-wave rectified. It is then analysed at coarse spatial resolution by a first-order mechanism identical to that used in the first-order channel. The first-order channel is insensitive to second-order motion, while the second-order channel responds to both first and second-order stimuli Fig. 4.

3.1. First-order motion processing

The initial stage of processing common to both channels is the removal of the mean luminance of the stimulus, which is assumed to occur at the retinal level in the human visual system (Shapley & Enroth-Cugell, 1984). This mean zero signal serves as the input to the fine scale motion detectors in the first-order channel. Prior to extracting coarse scale first-order motion the mean zero signal is low-pass filtered at progressively coarser spatial scales. The low-pass filters used are two-dimensional Gaussians in space and truncated exponentials in time. Filters of this form were chosen for their computational simplicity. Truncated exponentials are causal in time, and allow a simple recursive implementation (Fleet & Langley, 1995; Zanker, 1996; Clifford, Ibbotson & Langley, 1997). Filtering at coarser spatial scales is achieved by subsampling the image by a factor of two in each spatial

dimension between successive filtering operations.

Motion is extracted by a standard first-order motion energy computation (Adelson & Bergen, 1985) using Gabor filters in space and time. Gabor filters were chosen to provide a measure of the amount of motion information available from the stimulus, not necessarily as an accurate model of the motion analysis filters in the human visual system. While there is evidence of spatial filters with receptive fields resembling Gabor functions in the primary visual cortex (Pollen & Ronner, 1981), the visual system's temporal filters will necessarily be causal and are likely to be better described by the modulation of an asymmetric temporal envelope than by a Gabor function (Johnston & Clifford, 1995b).

3.2. Pre-processing in the second-order channel

Pre-processing in the second-order channel is similar to that for coarse scale motion in the first-order channel, except that the signal is full-wave rectified after the first low-pass filtering operation. The full-wave rectified signal is greater than or equal to zero everywhere. To avoid the effect of any DC response in the even-symmetric motion energy filters, either the signal mean must be removed or filters with no DC response such as high-order derivatives of Gaussians must be used rather than Gabors to obtain quadrature.

The pre-processing stage in the second-order channel acts as a texture grabber (Chubb & Sperling, 1991), converting variations in the spatial and temporal frequencies of the image signal into variations in the magnitude of the pre-processed signal. Such a scheme necessarily confounds spatial frequency and contrast by encoding their variations along a single dimension, but Werkhoven et al. (1993) provide psychophysical evidence that human perception of texture-defined motion is subject to the same confusion. Here we extend the spatial texture grabber proposed by Werkhoven et al. into the spatiotemporal domain by introducing a low-pass temporal filter. Encoding spatiotemporal frequency and contrast along a single dimension in this way, the model predicts that confusions will also exist between spatial frequency and temporal frequency, and between temporal frequency and contrast. Whether this is indeed the case in human texture-defined motion perception is an empirical question open to psychophysical investigation.

3.3. Computation of image velocity

Local motion energy computation in a given channel at a given scale is sufficient to provide an estimate of the direction of motion of the stimulus. Since all the psychophysical simulations presented here are two alternative forced choice direction discriminations, the model is implemented to this level. However, individual motion energy measurements do not contain the necessary information to support judgements of image velocity. The motion energy mechanism's constituent filters are band-pass in spatial and temporal frequencies. Thus the response of a given motion energy filter is a non-monotonic function of both spatial and temporal frequencies, as well as varying with image contrast, and does not depend in a simple way upon image velocity. The computation of image velocity from motion energy mechanisms requires the combination of outputs from filters at a range of spatial and temporal scales (Heeger, 1987), and possibly across channels. The question of whether speed computation involves the combination of information from the first and second-order channels is discussed below, but this stage of computation is not implemented in the modelling software. Full details of the implementation of the model are given in Appendix A.

4. Simulations and discussion

4.1. Motion from spatial frequency

The operation of the second-order channel on Green's (1986) motion from spatial frequency stimulus is shown in Fig. 6. Fig. 6A shows a frame from the stimulus

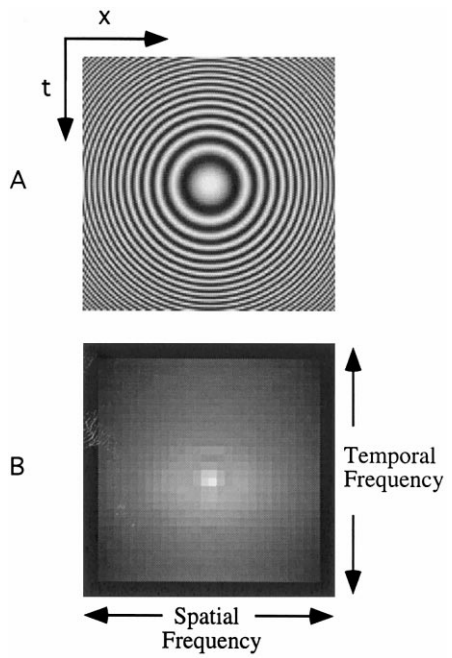

Fig. 5. Illustration of the operation of the texture grabber processing the stimulus prior to motion analysis in the second-order channel of the model. (A) A space-time slice through an artificial 'stimulus' input to the model. Spatial frequency increases linearly with spatial distance from origin of the space-time plot. Similarly, temporal frequency increases linearly with the distance in time. The response to this stimulus, shown in (B), thus represents the response of the pre-processing stage of the second-order channel as a function of spatial and temporal frequency. It will be seen that the texture grabber converts variations in spatial or temporal frequency in the stimulus to variations in the strength of the filtered signal, with the strongest signal at the lowest spatial and temporal frequencies.

sequence. Texture grabbing and further low-pass filtering transform the Gabor patches of differing spatial frequencies into blobs of different intensities (Fig. 6B). This pre-processed image sequence is then analysed by the motion energy mechanism. The motion energy mechanism is sensitive to the intensity differences between the blobs. When the Gabor patches in the stimulus rotate (Fig. 1) the motion between corresponding patches evokes a stronger response than the motion between dissimilar ones, effectively recovering the correct direction of rotation. Fig. 6C shows the result of computing motion energy in the horizontal direction following second-order pre-processing when the stimuli rotate clockwise. Lighter than the background denotes motion to the left, darker shows motion to the right, and the background gray-level represents no net motion. The response to the motion of the top patch is to some extent ambiguous: there is a strong rightward response along the trajectory of motion between corresponding patches, and a weaker leftward response between dissimilar patches. However, the net response for the top patch is one of rightwards motion. Similarly, the net response for the bottom patch is leftwards motion, indicating a clockwise rather than a counter-clockwise rotation.

Fig. 7 shows the psychophysical data from patient RA and a normal control on two conditions of the Boulton and Baker (1993) stimulus (from Vaina et al,. 1998, Fig. 5). The stimulus is a two-frame motion stimulus consisting of an array of Gabor patches on a mean luminance background. All the patches are displaced by the same amount between frames. The data are plotted as the percent error in a two alternative forced choice direction discrimination as a function of stimulus displacement in terms of the wavelength (lambda) of the carrier. The left-hand graphs show data obtained when there were 66 Gabor patches present on the screen, which Boulton and Baker term the 'dense' condition. On the right are data from the 'sparse' condition: 36 patches. The data from the control subject shown here replicate Boulton and Baker's finding that there is a qualitative difference in performance on the two conditions. Boulton and Baker argued that performance on the dense condition reflects the operation of a first-order motion channel, while a second-order channel mediates perception in the sparse case.

Fig. 8 shows the results of computational simulations with the fine scale first-order and coarse scale second-order channels on dense and sparse conditions of the Boulton and Baker stimulus (stimulus parameters for the simulation are given in Appendix B). The response of each channel is plotted as a direction index, calculated from the leftward and rightward motion energy responses (see Appendix B for details). Here, a direction index of $+1$ corresponds to motion energy entirely in the correct direction, -1 to motion in the incorrect direction, and zero to no motion or to equal and opposite

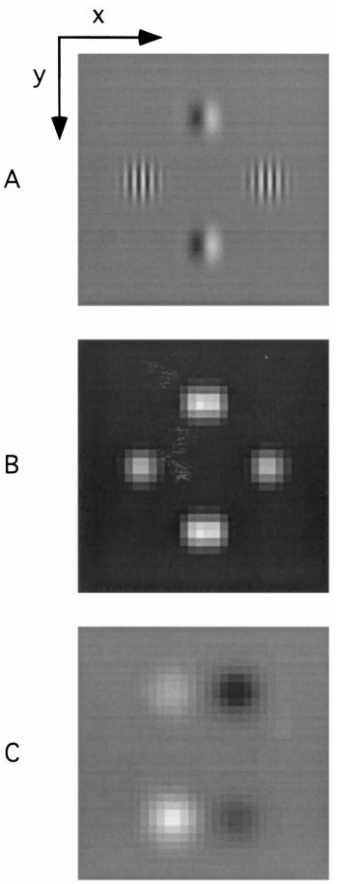

Fig. 6. (A) Single frame of the Green (1986) motion from spatial frequency stimulus (Fig. 1). Two pairs of Gabor patches describe a rotational motion on a mean luminance background. (B) A frame from the output of the texture grabber pre-processing the image in the second-order channel. The different spatial frequency Gabor patches produce responses of different magnitudes after texture grabbing, allowing the direction of the motion to be disambiguated. (C) A frame from the output of the motion energy detector in the second-order channel computing motion in the horizontal direction. The response to the motion of the top patch is to some extent ambiguous: there is a strong rightward response (white blob) along the trajectory of motion between corresponding patches, and a weaker leftward response (dark grey blob) between dissimilar patches. However, the net response for the top patch is one of rightwards motion. Similarly, the net response for the bottom patch is leftwards motion, indicating a clockwise rather than a counter-clockwise rotation.

leftward and rightward motion energies. The y-axis of the plots in Fig. 8A, B, D, E are inverted to facilitate comparison with the psychophysical data.

Fig. 8A, E shows that the results of the model are in accord with Boulton and Baker's conjecture. Performance of the control subject on the dense condition is closely predicted by the model's fine scale first-order channel (Fig. 7C and Fig. 8A), and on the sparse condition by the coarse scale second-order channel (Fig. 7D and Fig. 8E). However, the question remains of why the first-order channel should determine perceived direction in the dense condition and the second-order channel in the sparse. A possible answer lies in the relative

responsiveness of the first and second-order channels to the two stimulus conditions. As a measure of responsiveness we looked at the total motion energy (leftwards and rightwards) in each channel for each displacement in the two conditions. The results are shown in Fig. 8C, F, with the total motion energy of each channel normalised independently to the largest response found by that channel in either condition. The second-order channel is seen to respond more strongly to the sparse stimulus, while the largest responses of the first-order channel are seen with the dense stimulus condition.

Patient RA's data is similar to the control's on the sparse case, but markedly different on the dense case. Vaina et al. (1998) observe that the form of RA's data on the dense condition is similar to the form of his and the control's data on the sparse task, suggesting that RA's performance on both tasks is governed largely by second-order motion perception. The performance of the model's second-order channel on the dense stimulus condition is consistent with this interpretation (Fig. 8B). Both for RA and for the model's second-order channel the motion of the dense stimulus is rather ambiguous, but for both there is a tendency to respond correctly at shorter displacements and incorrectly at longer displacements. However, there is an interesting discrepancy between the predictions of the second-order channel of the model and RA's data on the dense stimulus condition: the direction index of the second-order channel's response is a monotonic function of displacement, while the graph of RA's data has a jagged appearance. It can be seen that the peaks and troughs in RA's data occur in phase with the peaks and troughs in both the data from the control subject and in the output of the model's first-order channel, suggesting that the graph of RA's data may be viewed as a superposition of a monotonically increasing second-order response and a weak first-order response. The model simulations in Fig. 8A–C suggest that we might expect such behaviour if RA had a weak residual sensitivity to first-order motion such that the responsiveness of his first-order channel to the dense stimulus was diminished to a level comparable to the second-order response.

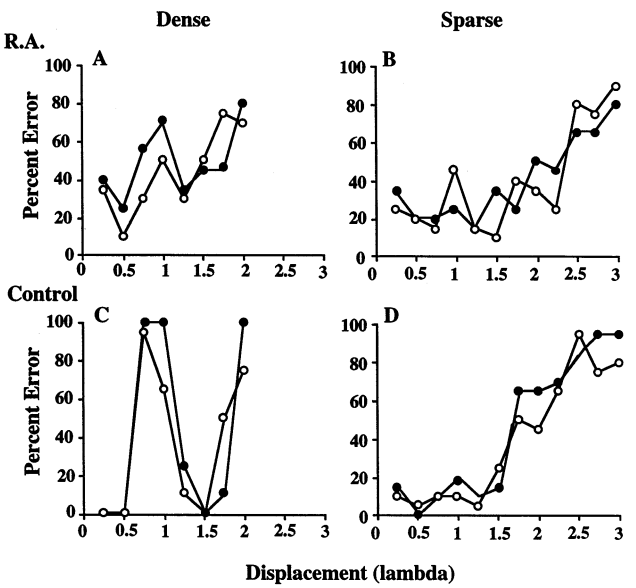

Fig. 7. The results of psychophysical direction discriminations on the Boulton and Baker (1993) Gabor array stimulus (Fig. 2). The top two panels show data from patient RA, the bottom two are from a normal control. The left-hand panels show results on the dense stimulus condition, the right-hand two on the sparse. Results are presented as the percentage of errors made in direction judgements as a function of stimulus displacement in the two-frame apparent motion sequence. The displacement is given in terms of the wavelength, lambda, of the carrier in the Gabor stimulus.

4.2. Motion from flicker

Fig. 9A shows a space-time plot of a flickering bar moving across a static background. The space-time plot is a slice normal to the y-axis through $I(x, y, t)$, the image sequence input to the model. The background is a random binary texture which does not change position from frame to frame. Between frames a variable proportion of the dots in a bar-shaped region of the image reverse their polarity. The flickering of the dots is the only stimulus attribute that defines the bar. The position of the bar changes smoothly over time during the image sequence, giving rise to oriented structure in the space-time plot and to second-order motion in the stimulus.

Fig. 9B is a space-time slice through the image sequence after second-order pre-processing. The low-pass texture grabber responds more strongly to the static background (zero temporal frequency) than to the flickering bar, so the bar appears in the pre-processed image as a dark bar on a light background. The difference in intensity between bar and background after second-order pre-processing allows the leftwards motion to be recovered by subsequent motion energy computation. Fig. 9C shows a space-time slice of the horizontal motion energy. Again, lighter than the background denotes leftwards motion, darker shows rightwards, and the background grey-level denotes static or no net horizontal motion. The fact no pixels darker than mid-grey appear in Fig. 9C indicates that no rightwards motion was detected in the image sequence. The leftwards motion response is seen to track the position of the bar over time.

4.3. Motion from contrast

The studies by both Plant and Nakayama (1993) and Greenlee and Smith (1997) used first and second-order grating stimuli. The first-order stimuli were sinusoidal luminance gratings. In the latter study these were superimposed upon a random binary texture (Fig. 10A). Fig. 10B shows that the model's first-order channel is able to recover the direction of motion of luminance gratings

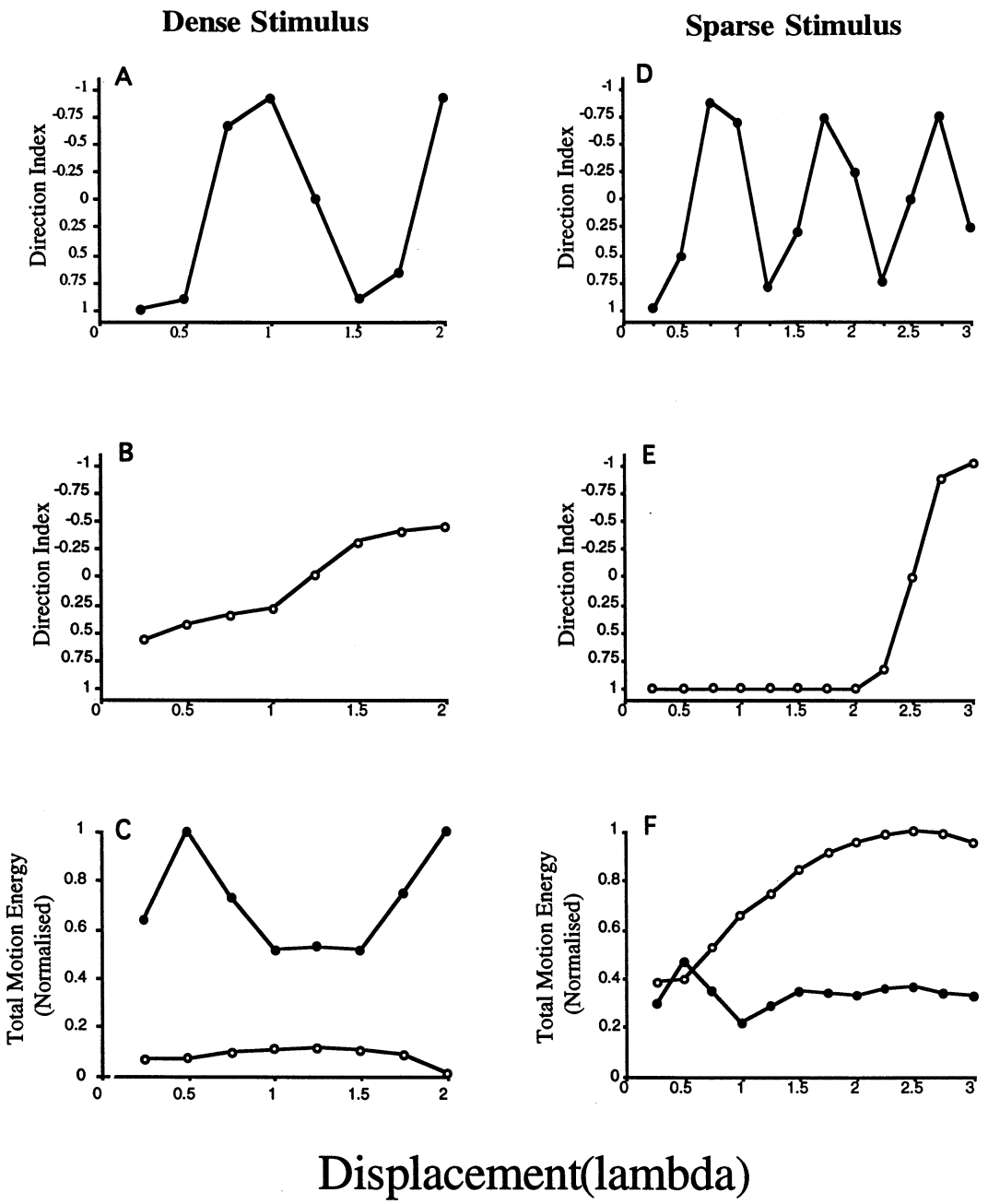

Fig. 8. Results of the model in simulations using a Gabor array stimulus (Fig. 2). Data for the dense condition are shown for (A) the first-order channel, and (B) the second-order channel of the model. (C) The activity of the first-order channel (solid symbols) and second-order channel (open symbols) in response to the dense stimulus. Data for the sparse condition are shown for (D) the first-order channel, and (E) the second-order channel of the model. (F) The activity of the first-order channel (solid symbols) and second-order channel (open symbols) in response to the dense stimulus.

with little effect from the static noise background. The convention followed in Fig. 10B, C is that mid-grey represents static motion or no net motion energy. Intensities lighter than this represent motion to the left, darker to the right. The fact that the darkest pixels in Fig. 10B are mid-grey shows that all the motion recovered from the image sequence is leftwards. Fig. 10C shows the output of the model's second-order channel to the same stimulus. Here, the motion energy measurements are random fluctuations about zero introduced by the presence of the static noise background.

Fig. 11 shows the output of the model to a second-order grating as used by Greenlee and Smith (1997). The stimulus in Fig. 11A is similar in appearance to the first-order stimulus in Fig. 10A. However, in the second-order stimulus the sinusoidal grating modulates the contrast of the background, while in the first-order stimulus a sinusoidal variation in luminance is added to

the background. The pre-filtering and rectification carried out in the model's second-order channel make its response a reliable indicator of the direction of motion of the contrast modulation (Fig. 11C). The first-order channel is insensitive to the motion of the contrast-defined grating (Fig. 11B).

Disruption of the model's second-order pathway does not have a severe effect on judgements of the velocity of first-order motion, as observed with Case 1. The model's second-order pathway is generally sensitive to first-order motions as well, although as implemented the texture grabber returns a spatiotemporally uniform field when presented with a sinusoidal luminance grating as used by Plant and Nakayama (1993). If information from the two pathways were pooled in the calculation of velocity then performance on some first-order speed and temporal frequency discriminations would be slightly impaired by disruption of the second-order pathway, although not for sinusoidal grating stimuli. In any case, the response of the first-order pathway would still allow the discriminations to be carried out accurately, whereas Case 1's Weber fraction for speed discrimination is elevated by a factor of three across a range of contrasts. We thus conclude that Case 1's performance is not consistent with disruption solely to a second-order motion mechanism. If Case 1's lesion affected an area containing predominantly velocity-tuned cells, but spared regions afferent to this containing direction-selective cells, then we would expect his performance on speed discrimination tasks to be severely impaired while his detection and DOM thresholds remained largely unaffected. Plant et al. (1993) suggest that Case 1's motion perception deficits are due to damage to the human homologue of MT.

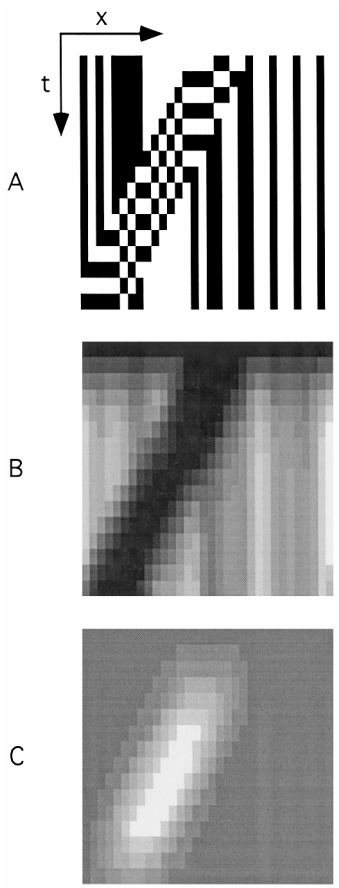

Fig. 9. (A) Space-time plot of the flickering bar stimulus (Fig. 3) used in simulations with the model. The trajectory of the flickering bar appears as an oriented structure against the static background (vertical structure). (B) The result of texture grabbing in the second-order channel, again represented as a space-time plot. The low-pass texture grabber responds more strongly to the static background than to the flickering bar, so after texture grabbing the bar appears as a dark region. (C) A space-time slice through the motion energy response of the model. Leftwards motion (brighter than the background) is recovered in the region defining the bar.

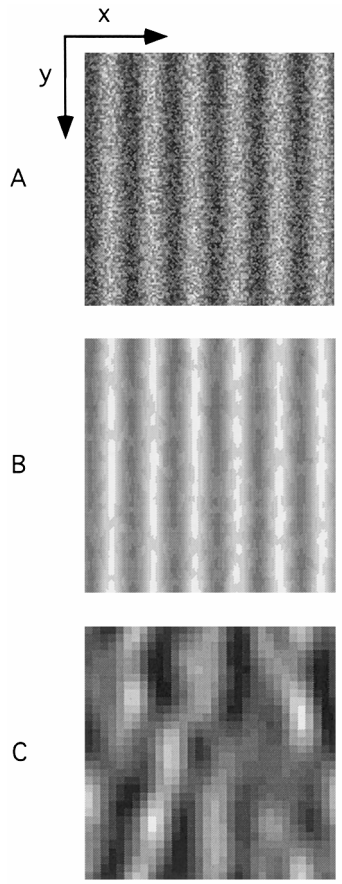

Fig. 10. (A) A frame from a first-order grating stimulus of the form used by Greenlee and Smith (1997). The stimulus consists of a luminance grating drifting across a static noise background. (B) The model's first-order channel is able to recover the direction of motion of luminance gratings with little effect from the static noise background. The convention followed is that mid-grey pixels represent static motion or no net motion energy. Intensities lighter than this represent motion to the left, darker to the right. The fact that all the pixel intensities are between mid-grey and white shows that all the motion recovered from the image sequence is leftwards. (C) The output of the model's second-order channel to the same stimulus. Here, the motion energy measurements are random fluctuations about zero introduced by the presence of the static noise background.

This is an area in which cells tuned to stimulus velocity have been reported (Rodman & Albright, 1987). Greenlee and Smith (1997) also reported a significant correlation between patients' thresholds on first and second-order speed discriminations. Their data, along with Case 1's, suggest either first and second-order motion information is combined in the computation of image velocity, or independent computations are carried out in a single region or adjoining regions.

Beat stimuli of the form used by Plant and Nakayama (1993) are not well suited to isolating second-order motion processing. Chubb and Sperling (1988) suggest that stimuli for studying second-order motion should be 'drift-balanced', so as not to contain systematic first-order motion components. They define a stimulus to be drift-balanced if its power spectrum in the frequency domain is symmetric with respect to temporal frequency. Beat stimuli are formed by the additive superposition of two sinusoidal components. A detailed analysis of stimuli of this kind is given by Fleet and Langley (1994). The frequency of what we perceive as the carrier of the beat stimulus is actually the mean

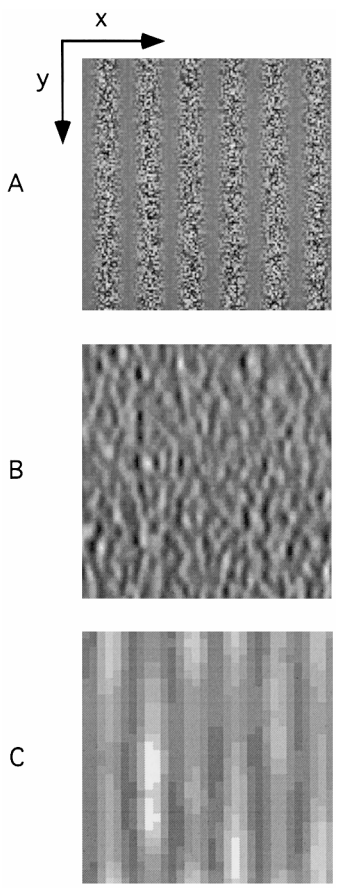

Fig. 11. (A) A frame from a second-order grating stimulus of the form used by Greenlee and Smith (1997). The stimulus consists of a drifting sinusoidal contrast envelope modulating a static noise background. (B) The model's first-order channel is unable to recover the motion of the stimulus. (C) The second-order channel correctly signals the direction of stimulus motion.

of the frequencies of the two components. When the carrier is static, as in the Plant and Nakayama experiment, the two components have equal and opposite temporal frequencies but differ in spatial frequency. Thus, the beat stimulus is not drift-balanced. The difference in spatial frequency of the components means that the stimulus will excite one member of an opponent pair of energy detectors more than the other, giving rise to a net directional response (Fig. 12A). When the spatial frequency of the carrier is less than the peak spatial frequency tuning of the motion energy filters the net response will be in the direction of motion of the beat. When the spatial frequency of the carrier is greater than that preferred by the filters, the net motion energy response will be in the opposite direction. In any case, unless the spatial frequency of the carrier exactly matches the spatial frequency tuning of motion energy filters, the beat stimulus will give rise to a consistent directional response which could form the basis of a direction discrimination judgement (Fig. 12B). The stimulus used by Plant and Nakayama had a carrier frequency of 0.5 cd, suggesting that all but the lowest spatial frequency channels would be able to recover the direction of motion of the beat correctly through first-order motion analysis.

Fig. 13 shows the results of simulations using first and second-order random-dot stimuli (Vaina & Cowey, 1996; Vaina et al., 1998). The direction index of the response of the first and second-order channels is plotted as a function of the coherence of the stimulus. Coarse spatial resolution was used to generate results from the first-order channel because the dynamic noise in the stimulus interferes with the recovery of motion at fine scales. Fig. 13A shows that the direction index of the response of the first-order channel increases roughly linearly with coherence of the first-order stimulus (solid symbols) while it fluctuates around zero for the second-order stimulus (open symbols). Thus the first-order channel is sensitive to the first-order motion, but is effectively blind to the second-order. Fig. 13B shows that the second-order channel responds to both the first and the second-order stimuli, but is more sensitive to the second-order.

From Fig. 13, we see that disruption of the second-order channel will severely impair performance on the second-order stimulus, as found with FD (Vaina & Cowey, 1996). When the first-order channel is disrupted, the model retains some residual sensitivity to first-order motion through its second-order channel. The second-order channel's sensitivity to first-order motion is lower than to second-order, so disruption of the model's first-order channel will cause coherence thresholds for first-order motion to rise above those for second-order. This is consistent with the data from RA. His coherence threshold for first-order motion in the hemifield contralateral to his lesion was elevated to

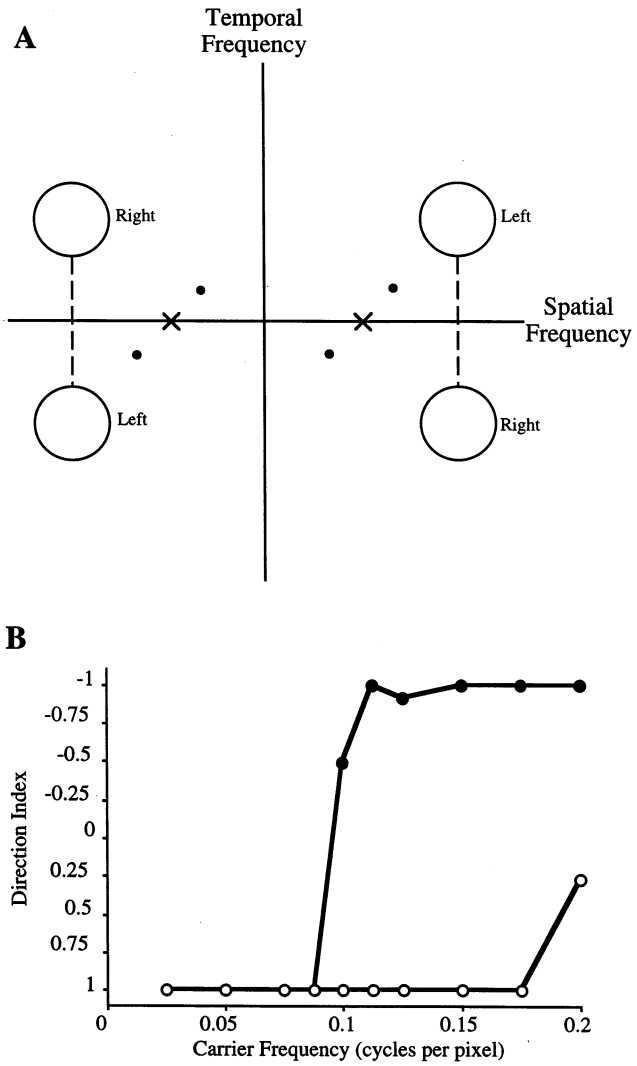

Fig. 12. (A) Schematic frequency space representation of the operation of motion energy filters on a contrast beat stimulus of the form used by Plant and Nakayama (1993). The positions of the dots correspond to the two sinusoidal components in the beat stimulus. The crosses are positioned on the spatial frequency axis at the spatial frequency of the carrier. The fact that the crosses are on the axis tells us that the carrier is static. Note that there is no power in the stimulus at the frequency of the carrier. The large open circles represent the preferred spatiotemporal frequencies of leftward and rightward motion energy filters. Both are band-pass in spatial and temporal frequency, preferring the same spatial frequencies but equal and opposite temporal frequencies. The frequency components of the stimulus fall closer to the preferred frequency of the leftwards filter than the rightwards filter. This causes the leftwards filter to respond more strongly than the rightwards when presented with the beat stimulus, giving rise to a net leftwards motion response. When the spatial frequency of the carrier is lower than the preferred spatial frequency of the motion energy filters, motion will be signalled in the direction of motion of the beat.(B) Direction indices of the responses of first-order (solid symbols) and second-order (open symbols) channels to beat stimuli with a range of carrier frequencies. The preferred spatial frequency tuning of the filters in the first-order channel is 0.1 cycles per pixel, and 0.025 in the second. Increasing the carrier frequency above the filters' preferred value reverses the direction of motion signalled by the first-order channel, but still gives a strongly directional signal.

about twice the level of his second-order motion coherence threshold (Vaina et al., 1998).

5. General discussion

We propose a simple computational model which is sensitive to motion defined by first-order variations in luminance, and to second-order motion defined by contrast, spatial frequency and flicker. The model consists of two parallel channels geared towards the perception of first and second-order motion respectively. The image signal in the second-order channel is processed by a texture grabbing mechanism before being subject to motion analysis. The mechanism of motion analysis used in the two-channels is identical, although the second-order channel operates only at a coarse spatial scale.

Fig. 14 shows schematically how we believe the computations carried out by the visual system might provide the basis for psychophysical judgements. At the lower levels of processing two parallel channels underlie the detection and direction of motion discrimination of first and second-order motion. Second-order motion is detected only in the second-order channel, while first-order motion is detectable by either the first or the second-order channel. Whether information from the first and second-order channels can be combined in detection and DOM tasks involving first-order stimuli is not clear from the patient data. The question of whether first and second-order motion pathways do interact in direction discrimination has been addressed using normal subjects by Edwards and Badcock (1995) and Nam and Chubb (1997). Using a random-dot stimulus, Edwards and Badcock showed that the presence of first-order 'noise dots' impaired the perception of global motion in a second-order stimulus, while second-order noise dots had no effect on the perception of first-order global motion. From this they concluded that the first and second-order motion channels do not interact in the discrimination of the direction of global motion, although the second-order channel is also sensitive to first-order motion. By investigating the perceived direction of motion of squarewave modulations of luminance and/or contrast, Nam and Chubb also found that the second-order pathway is highly sensitive to first-order motion while the first-order pathway is insensitive to second-order motion. However, any interaction between channels in these experiments would have reduced performance in direction discrimination on the first-order task, since the second-order channel is sensitive to the directional noise in the stimulus while the first-order channel is not. The fact that information from the second-order channel can be ignored when it is uninformative does not necessarily imply that information from the two channels cannot be combined when it is beneficial to do so.

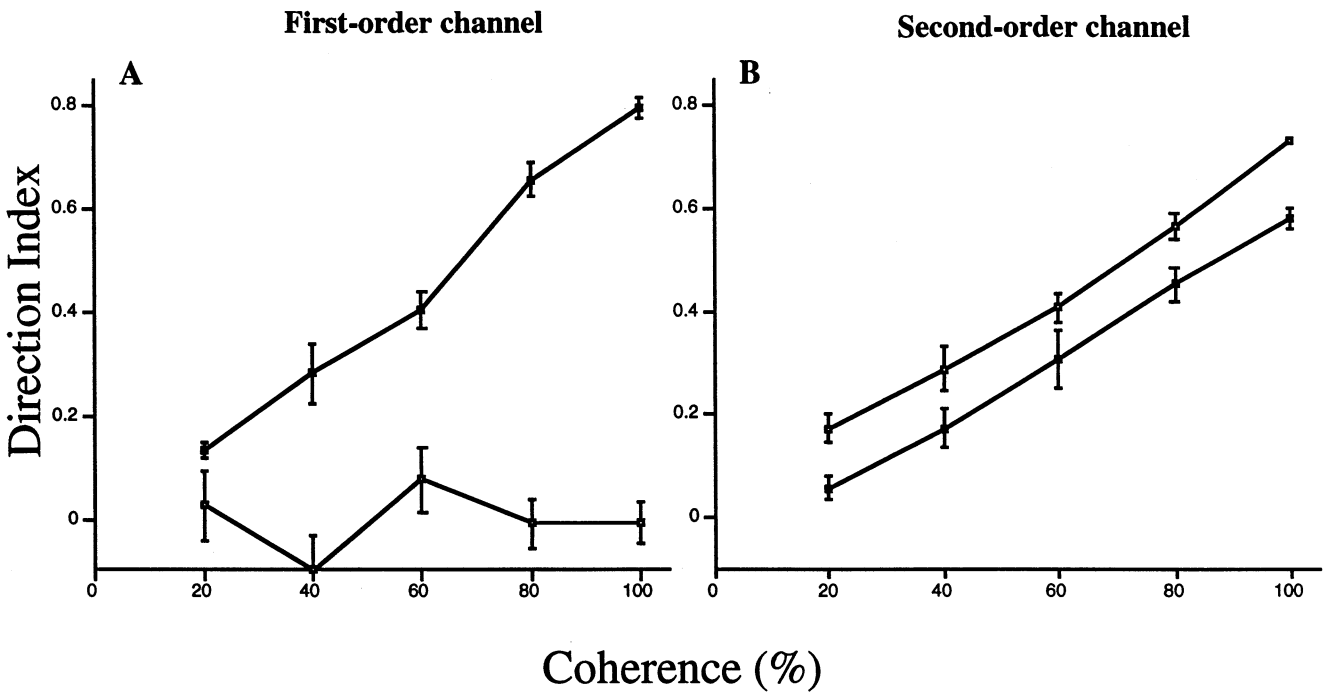

Fig. 13. Response of (A) the first-order channel, and (B) the second-order channel to first-order (solid symbols) and second-order (open symbols) random-dot stimuli. Each data point is the average of ten model simulations. Error bars are standard errors. Data is plotted in terms of the direction index of each channel's response as a function of the percentage coherence of the stimulus.

A functional architecture very similar to the one we propose here has previously been put forward to account for the results of motion after-effect studies using a flickering test stimulus (Nishida, Ashida & Sato, 1994; Nishida & Sato, 1995) . Nishida and co-workers found that only adaptation to first-order motion is able to produce significant motion after-effects in a subsequent static test stimulus (static MAE), but that both first and second-order adapting stimuli generate motion after-effects in a counterphase flickering test (flicker MAE). They interpreted their results as evidence that static MAE reflects the activity of a low-level system (the motion detectors in the first-order pathway) while flicker MAE reveals the behaviour of a high-level system (the motion detectors in the second-order pathway and the stage of first and second-order integration).

The computational and psychophysical hierarchy in Fig. 14 treats the calculation of image velocity as a higher-order process than the recovery of the direction of motion. In computational terms, this reflects the fact that direction of motion can be recovered from the calculation of motion energies at a single spatial and temporal scale while velocity computation requires the combination of information across scales (Heeger, 1987). Verghese and Stone (1996) have shown psychophysically that perceived speed depends upon the way in which an image is segmented, suggesting that it is mediated by higher-level processing mechanisms than the direction of motion.

Recent brain imaging studies have identified cortical regions located in V2 or V3 which respond more strongly to motion defined by contrast or flicker (Smith et al., 1997) than to first-order motion, and may be involved in the perception of motion boundaries (Dupont et al., 1997; Reppas et al., 1997). Following Wilson et al. (1992), we speculate that the neural substrate of the first-order channel is the direct projection from V1 to MT, while the second-order channel corresponds to the projection from V1 to MT via V2 and/or V3 (Fig. 14).

While the evidence from the recent neurological studies examined here strongly suggests the existence of two motion processing channels in the human visual system, one issue not addressed in this paper is whether there might be more than two channels. Cavanagh (1992) draws a distinction between automatic motion processing, which signals motion in the absence of attention, and motion processing mediated by the attentional tracking of image features. Lu and Sperling (Lu & Sperling, 1995a,b) suggest that such an attentionally modulated feature tracking system operates in addition to first and second-order motion channels. As with patient studies in second-order motion, we hope that future studies of attentional tracking in neurological subjects will help clarify the role of attention in motion perception.

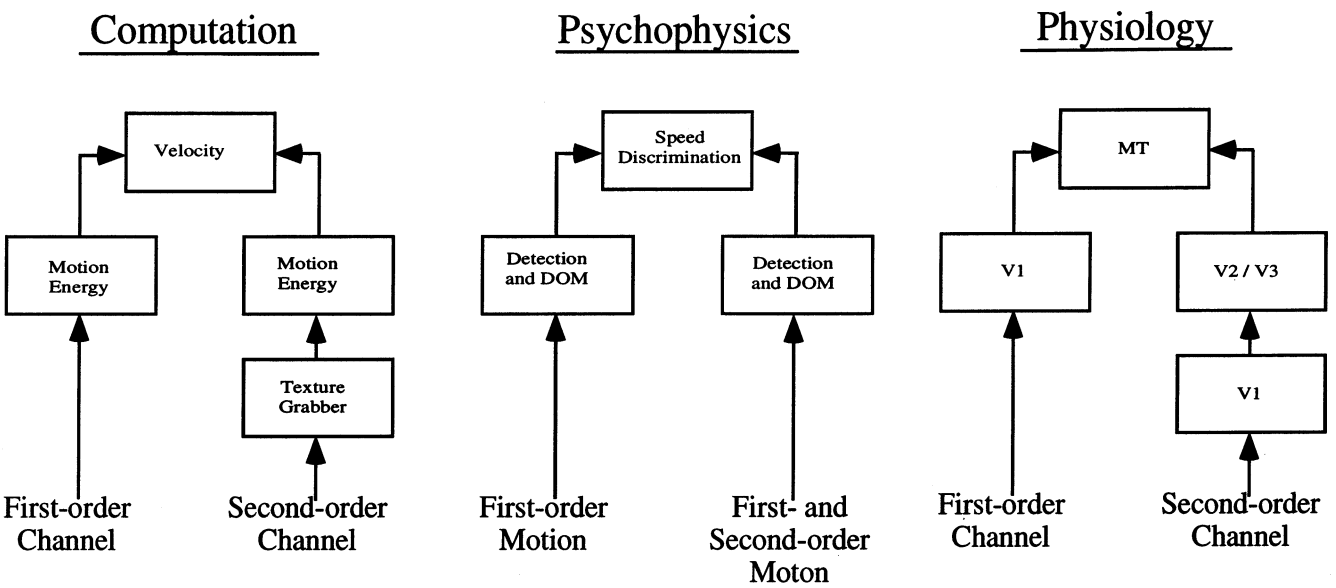

Fig. 14. Schematic diagram of the putative functional architecture of human motion processing in terms of the computations involved, the psychophysical decisions they support, and their probable neural substrates.

Acknowledgements

This work was supported by NIH Grant EY-2ROI-0781-06 to LMV. We are grateful to two anonymous reviewers for their thorough and constructive comments.

Appendix A. Implementation of model

The initial pre-processing stage common to both first and second-order channels is the removal of the mean luminance, or DC component, of the stimulus. The mean is calculated over the entire image on a frame-by-frame basis. The signal in the fine scale first-order channel is then operated upon directly by the motion energy detectors (Turano, 1991).

A.1. Coarse scale motion computation

For coarse scale motion, the signal is low-pass filtered in space and time. The filter used is space-time separable. The spatial filter is a two-dimensional Gaussian, $G(x, y)$:

$$G(x,y) = \frac{1}{2\pi\sigma^2} \cdot e^{-(x^2+y^2/2\sigma^2)}. \quad (1)$$

The spatial extent of the filter is governed by the parameter, σ, which here is set to 0.8 pixels. The two-dimensional Gaussian is itself separable into two one-dimensional Gaussians, so spatial filtering is carried out by successive convolution with two 3×1 masks. Spatial convolution wrapped around at the edge of images. The temporal filter, $E(t)$, is a first-order low-pass filter with an impulse response which decays exponentially over time:

$$E(t) = \frac{1}{\tau} \cdot e^{-t/\tau}, \qquad t \geq 0 \quad (2)$$

$$E(t) = 0, \; t < 0, \quad (3)$$

where τ is the time constant. The value of τ is necessarily positive, and gives a measure of the duration of temporal support. Here, τ is set to 0.5 frames. For efficiency, temporal filtering is implemented recursively using the following recursion relation:

$$O(t) = \beta \cdot O(T-1) + (1-\beta) \cdot I(T), \quad (4)$$

where T represents time sampled discretely in frames, $I(T)$ is the input to the filter at time T, $O(T)$ is the output of the filter, and β is a constant given by:

$$\beta = e^{-1/\tau}. \quad (5)$$

(For a derivation of the recursion relation see Clifford et al., 1997).

In the second-order channel the signal is then full-wave rectified and the mean removed. As above, the mean is calculated over the entire image on a frame-by-frame basis. The rectification and mean removal steps are the only stages at which coarse scale first and second-order pre-processing differ. The signal is then subsampled by a factor of two in both spatial dimensions, and spatial and temporal filtering repeated with the same temporal filters. The fact that the signal has been spatially subsampled means that convolution with the same filter kernels as previously represents spatial filtering at a coarser scale. The signal is then spatially subsampled and low-pass filtered in space and time once more before being input to the motion energy detectors.

A.2. Motion energy detection

Motion energies are computed from the outputs of quadrature pairs of filters oriented in space-time. Here, sine and cosine phase Gabor filters are used to approximate quadrature. The initial stage of filtering involves spatiotemporally separable filters, with space-time kernels, $\Psi(x,y,t)$, given by:

$$\Psi_{(x,y,t)} = \left(\frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_t}\right) \times \cdot e^{-1/2((x^2/\sigma_x^2)+(y^2/\sigma_y^2)+(t^2/\sigma_t^2))} \cdot \cos(k_x x + \phi_x) \times \cdot \cos(\omega t + \phi_t), \quad (6)$$

where σ_x, σ_y and σ_t define the extent of the three-dimensional Gaussian envelope, k_x and ω are the horizontal, vertical and temporal carrier frequencies, and ϕ_x and ϕ_t define the phase of the carrier relative to the envelope. Here, the preferred spatial and temporal frequencies, k_x and ω, of the filters are set to 0.1 cycles per pixel and 0.1 cycles per frame respectively. The values of σ_x, σ_y are 2.5 cycles per pixel, and σ_t is 2.5 cycles per frame, implemented in 15×1 convolution masks. The phase, ϕ, is set to either zero or $\pi/2$ to give the cosine and sine members of each quadrature pair. The outputs of these filters are then combined using trigonometrical identities to give space-time oriented filters (see Adelson & Bergen, 1985).

Appendix B. Details of simulations

Stimuli for the model simulations are sequences of 128×128 images of 16 frames duration. Leftwards and rightwards motion energies are calculated for each point in the pre-processed image sequence. When space-space (Fig. 6C, Fig. 10C, Fig. 11C) or space-time (Fig. 9C) slices through the motion energy output of the model are shown, the intensity of each pixel corresponds to the net horizontal motion energy at a point in space and time. The motion energy plotted is simply the signed difference between leftward and rightward motion energies, scaled to utilise the full brightness range.

B.1. Direction index

The direction index is calculated as a summary statistic of the directional responses of the horizontal motion detectors. The direction index is used to facilitate comparison of the output of the model with the response of psychophysical observers. A direction index of $+1$ denotes horizontal motion unambiguously in the correct direction, -1 unambiguously in the opposite direction, and zero totally ambiguous motion or no net motion. To convert the responses of the model into a formal psychophysical decision would require additional assumptions about the neural substrate of perceptual decisions (see Britten, Shadlen, Newsome & Movshon, 1992).

Leftwards and rightwards motion energies are calculated for each point in the ninth frame of the pre-processed image sequence. The ninth frame was chosen to minimise temporal edge effects from the beginning and end of the image sequence. A local directional response, $D(x, y, t)$, is calculated for each point from the leftwards and rightwards motion energies, $L(x, y, t)$ and $R(x, y, t)$, according to the following equation:

$$D(x,y,t) = \frac{L(x,y,t) - R(x,y,t)}{L(x,y,t) + R(x,y,t) + \alpha}, \quad (7)$$

where α is a constant given a small positive value to condition the quotient. The value of α is set at 10^{-7}. If α was given a large value then this would qualitatively affect the form of $D(x, y, t)$, as with low motion energies the directional response would be biased towards zero. However, in the simulations here the value of α was not manipulated as a parameter but fixed at a low value purely to prevent arithmetic exceptions in the execution of the model software.

To obtain a single direction index from the output of the model, the signed directional responses are simply summed over space and divided by the sum of their moduli to give a number between ± 1. This value is multiplied by the sign of the correct response (leftwards defined as positive) to convert from a left-right index into a correct-incorrect index:

$$\text{Direction index} = \text{sign (correct response)} \cdot \frac{\sum_{x,y} D(x,y,t)}{\sum_{x,y} |D(x,y,t)|}. \quad (8)$$

B.2. Stimulus parameters

B.2.1. Gabor array stimulus

The Gabor array stimulus used to obtain the results presented in Fig. 8 is constructed as follows. The first image in the two image sequence consists of three rows of Gabor patterns, $P(x, y)$:

$$P(x,y) = \frac{1}{2\pi\sigma^2} \cdot e^{-(x^2+y^2/2\sigma^2)} \cdot \cos(k_x x). \quad (9)$$

The Gabor functions have a carrier wavelength, $\lambda = 2\pi/k_x$, of 8.0 pixels and the standard deviation, σ,of the Gaussian envelope is 6.0 pixels. The centre-to-centre spacing of the Gabor patterns is 20.0 pixels in the dense condition, and 40.0 in the sparse. These values are in direct proportion to those used by Boulton and Baker (1993) and Vaina et al. (1998), who used a wavelength of 16 pixels and σ of 12 pixels. The first image is present for eight frames. The second image is simply a shifted version of the first, with wraparound at the edges, again presented for eight frames.

B.3. Random-dot patterns

The random-dot patterns used to obtain the results presented in Fig. 13 are constructed as follows. The dynamic noise background has a contrast of 0.2, and each noise block is a 2×2 square of pixels. The polarity of each noise block varies randomly from frame to frame. Each dot in the dot pattern is an 8×8 square of pixels. The density of the dots is 8%. In the first-order pattern, the mean luminance of the dots is 1.3 times the mean luminance of the background, while the contrast is 0.2, the same as for the background. In the second-order pattern, the mean luminance of the dots is the same as the mean luminance of the background, while the contrast is 0.6. These values are the same as those used by Vaina et al. (1998) with a mean luminance of 9.5 Cdm^{-2}. At each frame the identity of each dot is randomly assigned to be 'signal' or 'noise', according to the coherence level of the stimulus. Signal dots all translate in the same direction at a rate of one pixel per frame, while noise dots are randomly re-positioned within the image so as not to coincide with another dot.

References

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

Albright, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science*, *255*, 1141–1143.

Benton, C. P., Johnston, A., & McOwan, P. W. (1997). Perception of motion direction in luminance and contrast-defined reversed-phi motion sequences. *Vision Research*, *37*, 2381–2399.

Boulton, J. B., & Baker, C. L. (1991). Motion detection is dependent on spatial I frequency not size. *Vision Research*, *31*, 77–87.

Boulton, J. B., & Baker, C. L. (1993). Different parameters control motion perception above and below a critical density. *Vision Research*, *33*, 1803–1811.

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

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, *12*, 4745–4765.

Cavanagh, P. (1992). Attention-based motion perception. *Science*, *257*, 1563–1565.

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.

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

Chubb, C., & Sperling, G. (1991). Texture quilts: basic tools for studying motion-from-texture. *Journal of Mathematical Psychology*, *35*(4), 411–442.

Clifford, C. W. G., Ibbotson, M. R., & Langley, K. (1997). An adaptive Reichardt detector model of motion adaptation in insects and mammals. *Vision Neuroscience*, *14*, 741–749.

Cropper, S. J. (1994). Velocity discrimination in chromatic gratings and beats. *Vision Research*, *34*, 41–48.

Derrington, A. M., & Badcock, D. R. (1985). Separate 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.

Dupont, P., De Bruyn, B., & Vandenberghe, R. (1997). The kinetic occipital region in human visual cortex. *Cerebral Cortex*, *7*, 283–292.

Dosher, B. A., Landy, M. S., & Sperling, G. (1989). Kinetic depth effect and optic flow-I. 3D shape from Fourier motion. *Vision Research*, *29*, 1789–1813.

Edwards, M., & Badcock, D. R. (1995). Global motion perception: no interaction between the first and second order motion pathways. *Vision Research*, *35*, 2589–2602.

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

Fleet, D. J., & Langley, K. (1995). Recursive filters for optical flow. *IEEE Transactions*, *17*, 61–67.

Geesaman, B. J., & Andersen, R. A. (1996). The analysis of complex patterns by form/cue invarient MSTd neurons. *Journal of Neuroscience*, *16*(15), 4716–4732.

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

Gegenfurtner, K. R., Kiper, D. C., & Levitt, J. B. (1997). Functional properties of neurons in macaque area V3. *Journal of Neurophysiology*, *77*, 1906–1923.

Gorea, A. (1995). Spatiotemporal characterization of a Fourier and non-Fourier motion system. *Vision Research*, *35*, 907–914.

Greenlee, M. W., & Smith, A. T. (1997). Detection and discrimination of first and second-order motion in patients with unilateral brain damage. *Journal of Neuroscience*, *17*, 804–818.

Green, M. (1986). What determines correspondence strength in apparent motion? *Vision Research*, *26*, 599–607.

Grzywacz, N. M., Watamaniuk, N. J., & McKee, S. P. (1995). Temporal coherence theory for the detection and measurement of visual motion. *Vision Research*, *35*, 3183–3203.

Harris, L. R., & Smith, A. T. (1992). Motion defined by second-order characteristics does not evoke optokinetic nystaimus. *Vision Neuroscience*, *9*, 565–570.

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

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

Johnston, A., & Clifford, C. W. G. (1995a). Perceived motion of contrast modulated gratings: predictions of the multi-channel gradient model and the role of full-wave rectification. *Vision Research*, *35*, 1771–1783.

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

Johnston, A., & Benton, C. P. (1997). Speed discrimination thresholds for first and second-order bars and edges. *Vision Research*, *37*, 2217–2226.

Landy, M. S., Dosher, B. A., Sperling, G., & Perkins, M. E. (1991). The kinetic depth effect and optic flow-II. First and second-order motion. *Vision Research*, *31*, 859–876.

Ledgeway, T. (1994). Adaptation to second-order motion results in a motion after-effect for directionally-ambiguous test stimuli. *Vision Research*, *34*, 2879–2889.

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.

Ledgeway, T., & Smith, A. T. (1995). The perceived speed of second-order motion and its dependence on stimulus contrast. *Vision Research*, *35*, 1421–1434.

Ledgeway, T., & Smith, A. T. (1997). Changes in perceived speed following adaptation to first-order and second-order motion. *Vision Research*, *37*, 215–224.

Lu, Z. L., & Sperling, G. (1995a). Attention generated apparent motion. *Nature*, *377*, 237–239.

Lu, Z. L., & Sperling, G. (1995b). The functional architecture of human visual motion perception. *Vision Research*, *35*, 2697–2722.

Mather, G. (1991). First-order and second-order visual processes in the perception of motion and tilt. *Vision Research*, *31*, 161–167.

Mather, G., & West, S. (1993). Evidence for second-order motion detectors. *Vision Research*, *33*, 1109–1112.

Mather, G., & Tunley, H. (1995). Motion detection in interleaved randomdot patterns: evidence for a rectifying nonlinearity preceding motion analysis. *Vision Research*, *35*, 2117–2126.

McCarthy, J. E. (1993). Directional adaptation effects with contrast modulated stimuli. *Vision Research*, *33*, 2653–2662.

Nam, J. H., & Chubb, C. (1997). The mechanism that detects texture-defined motion is highly sensitive to luminance-defined motion. *Investigative Ophthalmology and Vision Science*, *38*, 237.

Nishida, S., & Sato, T. (1995). Motion after-effect with flickering test patterns reveals higher stages of motion processing. *Vision Research*, *35*, 477–490.

Nishida, S., & Sato, T. (1992). Positive motion after-effect induced by bandpass-filtered random-dot kinematograms. *Vision Research*, *32*, 1635–1646.

Nishida, S. (1993). Spatiotemporal properties of motion perception for random-check contrast modulations. *Vision Research*, *33*, 633–646.

Nishida, S., Ashida, H., & Sato, T. (1994). Complete interocular transfer of motion after-effect with flickering test. *Vision Research*, *34*, 2707–2716.

Nishida, S., Edwards, M., & Sato, T. (1997). Simultaneous motion contrast across space: involvement of second-order motion? *Vision Research*, *37*, 199–214.

O'Keefe, L. P., Carandini, M, Beusmans, J. M. H., & Movshon, J. A. (1993). MT neuronal responses to Ist and 2nd-order motion. *Society of Neuroscience* (*Abstr.*), *19*, 1283.

O'Keefe, L. P., & Movshon, J. A. (1996). First and second-order motion processing in the superior temporal sulcus of the alert macaque. *Society of Neuroscience* (*Abstr.*), *22*, 716.

O'Keefe, L. P., & Movshon, J. A. (1997). Neuronal responses to first and second-order motion in the superior temporal sulcus of the alert macaque. *Investigative Ophthalmology and Vision Science*, *38*, 238.

Pantle, A. (1992). Immobility of some second-order stimuli in human peripheral vision. *Journal of the Optical Society of America A*, *9*, 863–867.

Papathomas, T. V., Gorea, A., & Chubb, C. (1996). Precise assessment of the mean effective luminance of texture patches an approach based on reverse-phi motion. *Vision Research*, *36*, 3775–3784.

Petersik, J. T. (1995). A comparison of varieties of second-order motion. *Vision Research*, *35*, 507–517.

Plant, G. T., & Nakayama, K. (1993). The characteristics of residual motion perception in the hemifield contralateral to lateral occipital lesions in humans. *Brain*, *116*, 1337–1353.

Plant, G. T., Laxer, K. D., Barbaro, N. M., & Schiffman, J. S. (1993). Impared visual motion perception in the contralateral hemifield following unilateral postenor cerebral lesions in humans. *Brain*, *116*, 1303–1335.

Pollen, D. A., & Ronner, S. F. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science*, *212*, 1409–1411.

Reichardt, W. (1961). *Autocorrelation, a principle for the evaluation of sensory information by the central nervous system*. New York: Wiley.

Reppas, J. B., Niyogi, S., Dale, A. M., Sereno, M. I., & Tootell, R. B. H. (1997). Representation of motion boundaries in retinotopic human visual cortical areas. *Nature*, *388*, 175–179.

Rodman, H. R., & Albright, T. (1987). Coding of visual stimulus velocity in area MT of the macaque. *Vision Research*, *27*, 2035–2048.

Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Progress in Retina Research B*, *3*, 263–346.

Smith, A. T. (1994). Correspondance-based and energy-based detection of second-order motion in human vision. *Journal of the Optical Society A*, *11*, 1940–1948.

Smith, A. T., Hess, R. F., & Baker, C. L. (1994). Direction identification thresholds for second-order motion in central and peripheral vision. *Journal of the Optical Society of America A*, *11*, 506–513.

Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M., & Hennig, J. (1997). Second-order motion may be detected by visual area V2 in humans: an fMRI study. *Society of Neuroscience* (*Abstr.*), *23*, 2230.

Smith, A. T., & Ledgeway, T. (1997). Separate detection of moving luminance and contrast modulations: fact or artifact? *Vision Research*, *37*, 45–62.

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

Solomon, J. A., & Sperling, G. (1995). Ist and 2nd-order motion and texture resolution in central and peripheral vision. *Vision Research*, *35*, 59–64.

Stoner, G. R., & Albright, T. D. (1992). Motion coherency rules are form-cue invariant. *Vision Research*, *32*, 465–475.

Taub, E., Victor, J. D., & Conte, M. M. (1997). Nonlinear processing in shortrange motion. *Vision Research*, *37*, 1459–1477.

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

Turano, K. (1991). Evidence for a common motion mechanism of luminance-modulated and contrast-modulated patterns: selective adaptation. *Perception*, *20*, 455–466.

Vaina, L. M., & Cowey, A. (1996). Impairment of the perception of second order motion but not first order motion in a patient with unilateral focal brain damage. *Proceedings of the Royal Society of London B*, *263*, 1225–1232.

Vaina, L. M., Makris, N., Kennedy, D., & Cowey, A. (1998). The selective impairment of the perception of first-order motion by unilateral cortical brain damage. *Vision Neuroscience*, *5*, 333–347.

van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, *2*, 300–321.

Verghese, P., & Stone, L. S. (1996). Perceived visual speed constrained by image segmentation. *Nature*, *381*, 161–163.

Victor, J. D., & Conte, M. M. (1990). Motion mechanisms have only limited access to form information. *Vision Research*, *30*, 289–301.

Victor, J. D., & Conte, M. M. (1992). Coherence and transparency of moving plaids composed of Fourier and non-Fourier gratings. *Perception & Psychophysics*, *52*, 403–414.

Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: a single channel theory. *Vision Research*, *33*, 463–486.

Werkhoven, P., Sperling, G., & Chubb, C. (1994). Perception of apparent motion between dissimilar gratings: spatio-temporal properties. *Vision Research*, *34*, 2741–2759.

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

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

Zanker, J. M. (1996). On the elementary mechanism underlying secondary motion processing. *Philosophical Transactions of the Royal Society of London B*, *351*, 1725–1736.

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