Perception of Apparent Motion Between Dissimilar Gratings: Spatiotemporal Properties

PETER WERKHOVEN,*† GEORGE SPERLING,*‡ CHARLES CHUBB*§

What determines the strength of texture-defined apparent motion perception when the stimulus has no net directional energy in the Fourier domain? In a previous paper [Werkhoven, Sperling & Chubb (1993) *Vision Research, 33,* 463–485] we demonstrated the counterintuitive finding that the *correspondence* in spatial frequency and in modulation amplitude between neighboring patches of texture in a spatiotemporal motion path are irrelevant to motion strength. Instead, we found strong support for what we call a single channel or *one-dimensional* motion computation: a simple nonlinear transformation of the image, followed by standard motion analysis. Here, we further studied the dimensionality of the motion computation in a parameter space that includes texture orientation and stimulus display rate in addition to texture spatial frequency and modulation amplitude. We used ambiguous motion displays in which one motion path, consisting of patches of nonsimilar texture, competes with another motion path comprised entirely of similar texture patches. The data show that motion between dissimilar patches of texture that are orthogonally oriented, have a two octave difference in spatial frequency and differ 50% in modulation amplitude can easily dominate motion between similar patches of texture. A single channel accounts for more than 70% of texture-from-motion strength for the parameter space examined and this channel is invariant for stimulus display rates varying over a four-fold range.

Second-order motion   Motion energy   Spatial frequency   Orientation

INTRODUCTION

In a previous paper (Werkhoven, Sperling & Chubb, 1993), we examined the perception of apparent motion between dissimilar gratings and introduced a paradigm to determine the nature of the underlying motion computation. Before describing the goals of the present paper, we will summarize some of our previous results.

We used a motion competition scheme that contained a heterogeneous motion path between dissimilar gratings and a homogeneous motion path between similar gratings (Fig. 2). Subjects were asked to indicate the perceptually dominant motion path. The phase of the gratings was randomized within a motion path, and the gratings had mean luminance equal to that of the background. These conditions insured that the stimuli were *microbalanced* (Chubb & Sperling, 1988, 1991)— whatever motion they displayed was invisible to mechanisms that correlate stimulus luminance across space and

time in computing image motion. The simplest models to explain such texture-defined motion involve nonlinear preprocessing of the stimulus ("texture grabbing") followed by standard motion analysis.

Our motion competition scheme was novel in the sense that only one heterogeneous motion path competed with only one homogeneous motion path. By showing that a heterogeneous motion path (between grating patches of textures *s* and *v*) could easily dominate a homogeneous motion path (between grating patches purely of type *s*), we disproved the well-entrenched presumption that strength of texture-defined motion was greatest between patches of texture with similar properties. Previous experiments (e.g. Watson, 1986; Green, 1986) had seemed to support the view that texture-defined motion is determined by correspondence matching. However, the apparent domination of heterogeneous motion paths by homogeneous paths was shown to be inherent to the competition schemes used and therefore was inconclusive.

A single-channel motion computation

The findings of our previous experiments were explained in terms of a *single-channel motion computation.* That is, the motion computation consists of a single preprocessing stage followed by standard motion analysis (Fig. 1).

*Department of Psychology and Center for Neural Science, New York University, New York, NY 10003, U.S.A.

†Present address: TNO, Human Factor Research Institute, Soesterberg, Postbus 23, 376926, The Netherlands.

‡Present address: Department of Cognitive Science, SST, University of California at Irvine, Irvine, CA 92717, U.S.A.

§Present address: Department of Psychology, Rutgers University, New Brunswick, NJ 08903, U.S.A.

The preprocessing stage consists of a single linear spatial filter followed by rectification. We call this preprocessing transformation a *texture grabber*, and its output is called *activity*. Our previous results indicated that the activity of a patch of sinusoidal grating increases monotonically with the modulation amplitude of the grating and decreases monotonically with its spatial frequency. Motion strength was shown to be proportional to the product of the activities of two consecutive grating patches in a motion path.

The proposed motion computation (see Fig. 2) implies that the strength (S_{he}) of the heterogeneous path (the product of the activities of patches s and v) balances the strength (S_{ho}) of the homogeneous path (the product of the activities of patches s and s) if and only if the activity of v is equal to the activity of s. Yet, as we observed in our previous experiments, there exist drastically dissimilar gratings with equal activity. Thus, the similarity between the gratings in a motion path is by no means decisive in determining motion strength.

Let s and v be textures with equal activity. Under the single channel model sketched above, a stimulus (Scheme I, Fig. 2) that pits a homogeneous motion path comprised of patches of s against a heterogeneous path comprised of alternating patches of s and v will then elicit a balanced (perfectly ambiguous) global motion percept, i.e. the homogeneous and heterogeneous paths are of exactly equal perceptual strength. The same is true of a stimulus (Scheme II, Fig. 3) that pits a homogeneous motion path comprised of v patches against a heterogeneous path comprised of alternating v and s patches.

Conversely, we have shown theoretically that if it holds across all textures s, v that Scheme I is balanced only if Scheme II is balanced, then the data can be explained by a single channel motion model of the sort described above. We call this important empirical property *transition invariance*. A violation of transition invariance indicates a multi-channel motion computation. Therefore, the transition invariance test is a powerful tool to determine the dimensionality of the motion computation within a certain parameter space.
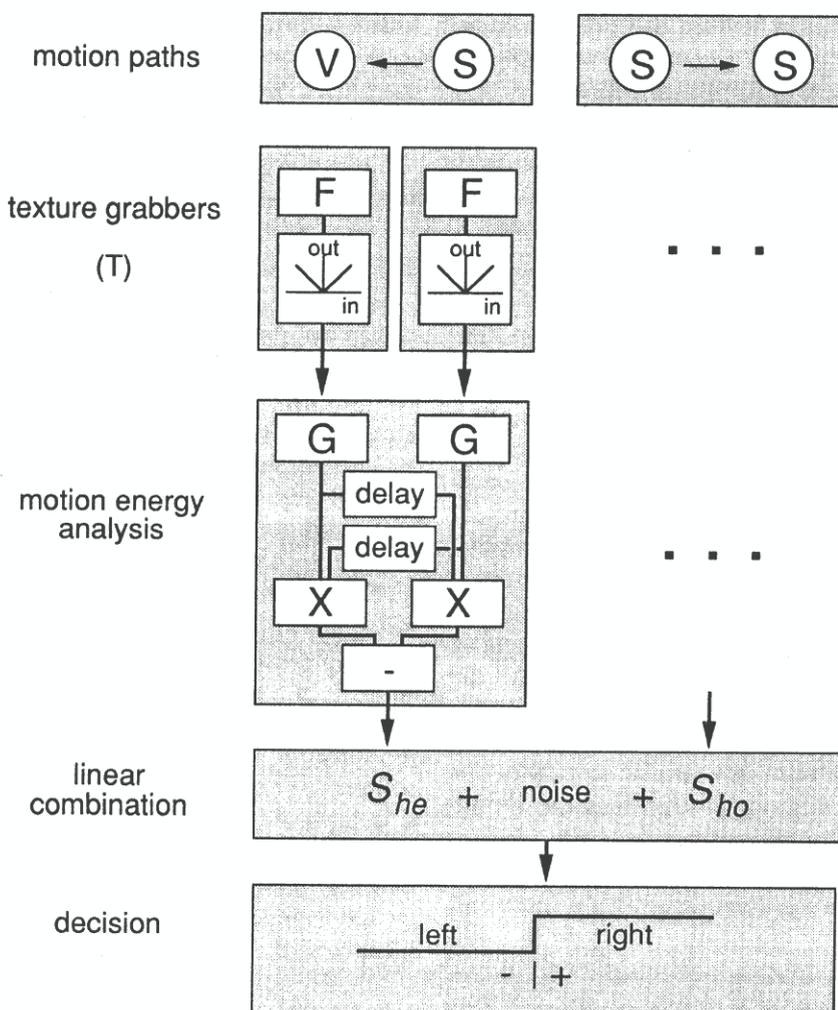

FIGURE 1. Diagram of a *single-channel* motion computation. First stimulus amplitude is extracted followed by a linear spatial filter F and rectification. The spatial filter together with the rectification is called 'texture grabber" (the first stage). The output of the texture grabber is called activity. The second stage (motion energy analysis) is basically a coincidence detector: it computes the product of the delayed activity (after linear spatial filtering by G) at location 1 with the current activity at location 2. Response variability across trials is due to internal noise. The heterogeneous path is dominant whenever the *net* motion strength $S_{he} - S_{ho}$ in the direction of the heterogeneous motion path (after adding noise) is positive (decision stage).

Aim of this study

In Werkhoven *et al.* (1993), we examined a two-dimensional parameter space: spatial frequency and modulation amplitude. We found that transition invariance held across the family of textures corresponding to points in this parameter space. For these parameters, texture-defined motion perception could be modeled by a single channel motion computation: a computation consisting of a single texture grabber followed by standard motion analysis. Here we ask whether transition invariance holds for a wider range of textural properties including grating *orientation* in addition to modulation amplitude and spatial frequency.

The relative orientation of gratings in a motion path was not varied in our previous experiments (the orientation of the gratings was kept orthogonal to the motion direction). Therefore, possible texture grabbers that are tuned to other orientations relative to the motion direction were not stimulated. If texture grabbers tuned to different orientations feed into separate motion analysis channels, the computation is essentially multi-dimensional. Hence, we would expect a violation of transition invariance. So, the first question is: is texture defined motion processed by separate motion channels that contain texture grabbers tuned to different orientations?

Secondly, the display rate of the motion stimulus was not varied in previous experiments. Frames were shown contiguously in time. Each frame was shown for 133 msec and all stimuli had a temporal period of four frames. Thus, stimulus display rate was 1.88 Hz. It is known that the temporal characteristics can affect the motion computation involved. For example, display rate is one of the properties of an apparent motion stimulus that has been asserted to distinguish short-range from long-range motion perception (Anstis, 1970; Pantle & Picciano, 1976). One might argue that, by increasing the duration of each frame, the processing time for each grating increases, thereby facilitating a more detailed comparison between gratings in a motion path. As a result, we might expect transition invariance to break down for lower display rates.

Here, we examine the effects of spatial frequency, modulation amplitude, orientation and display rate on the strength of apparent motion perception between dissimilar gratings.

METHOD

The strength of apparent motion perception was measured using two motion competition schemes (Schemes I and II). Both motion competition schemes contain a single homogeneous motion path (motion between identical textures) that competes with a single heterogeneous motion path (motion between different textures). It will be shown in the following sections that an analysis and comparison of the results of both

schemes allows us to distinguish between *single-channel* and *multi-channel* motion computations.

Stimulus

Motion Competition Scheme I. Motion Competition Scheme I consisted of a series of eight frames $(f_1, f_2, \ldots f_8)$ shown successively in time.

In Fig. 2, we show a sketch of the frames. The first frame (f_1) contained an annulus of patches of alternated texture type s and v at regular positions (see Fig. 2, left). The centers of the patches are on an annulus with a radius $\rho_0 = 1.56$ deg. The patches were regularly distributed around the annulus. Each patch had a circular shape which was the result of windowing a texture with a blurred circular aperture. The circular aperture had a diameter of 1.03 deg. The standard deviation of the Gaussian blurring kernel which was used to blur the aperture was 4.0 min are (see Fig. 4). A formal description of the textures seen through the apertures is given in the next section.

Frame f_2 was similar to frame f_1, except that patches of texture v were replaced by a uniform field of background luminance. Furthermore, f_2 was rotated around the center of the annulus over 22.5 deg with respect to frame 1 (see Fig. 2, left). In a sequence of frames, frame f_{n+2} was identical to frame f_n, except for a rotation around the center over 45 deg.

The frames were shown contiguously in time (no interval between frames). The presentation time of each frame ("frame display time") was $\tau = 67$, 133 or 267 msec in different experiments yielding temporal frequencies of 3.75, 1.88 and 0.94 Hz respectively. The presentation time of the sequence was $8\tau = 0.533$, 1.066 and 4.267 sec respectively. The displacement or distance separating the two nearest patches between frames was 36.5 min arc.

The ambiguous motion stimulus described above contains two motion paths. This can be understood most easily using a diagram in which we show the angular positions (ϕ) of the patches of texture for successive frames. Angular position is measured clockwise relative to the vertical. Such a diagram is shown in Fig. 2, right. Note that the rows of patches correspond to frames 1, 2, 3 and 4 respectively.

When frame f_n and frame f_{n+1} were presented in succession, two matches between patches of frame f_n and patches of frame f_{n+1} were likely *a priori*. The first match was a homogeneous match between patches of identical texture s rotated clockwise around the center of the annulus by 22.5 deg (indicated in the diagram by the arrow pointing down and to the right). The second match was a heterogeneous counter-clockwise match between patches of texture v and patches of texture s (indicated by the arrow pointing down and to the left). Matches between frames f_n and f_{n+2} are purely ambiguous. Other matches were only possible between patches of frames f_n and f_{n+3} over larger angles and across larger time intervals.

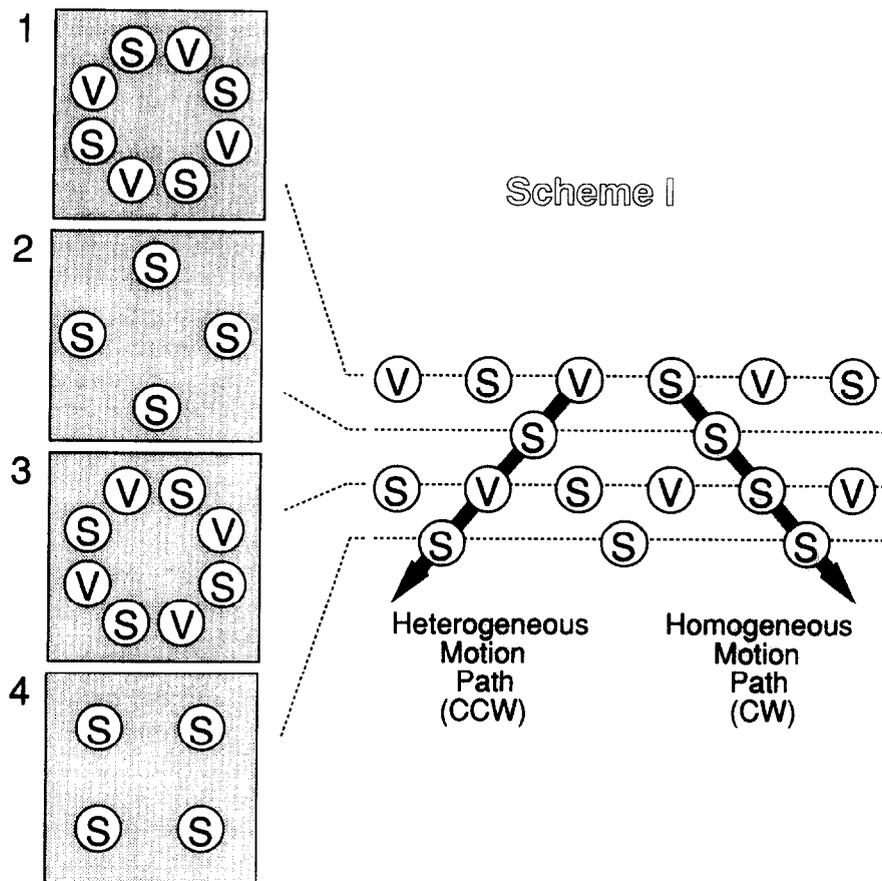These displays contain homogeneous and heterogeneous motion paths in opposite directions. By

FIGURE 2. Motion competition Scheme I. Left: a series of frames $(f_1, f_2, \ldots)$ was shown successively in time (for details see section). The first frame (f_1) contained an annulus of "disk shaped" patches of alternated texture type s and v at regular positions drawn against a uniform background. Angular position ϕ is measured clockwise with respect to the vertical. The second frame (f_2) is similar to frame f_1, except that the low frequent patches of texture v are now replaced by a uniform patch of background luminance. Furthermore, f_2 is rotated (clockwise) around the center of the annulus over an angle of 22.5 deg with respect to frame f_1. In a sequence of frames, frame f_{n+2} is identical to frame f_n, except for a rotation around the center over an angle of 45 deg (clockwise). Right: each row of patches results from cutting the corresponding annulus at the lowest point and stretching them. This schematic presentation reveals the competing motion paths more clearly. Angular positions ϕ are now along the horizontal axis. Time (or frame number) is along the vertical axis. When frame f_n and frame f_{n+1} are presented in succession, two motion paths are *a priori* likely. A homogeneous motion path: clockwise matches (CW) between patches of identical texture s (indicated by the arrow pointing down and right). A heterogeneous motion path: counter-clockwise (CCW) matches between patches of texture s and patches of texture v (indicated by the arrow pointing down and left).

randomizing the sign of the rotation, we randomize the respective directions of the two motion paths (which are, of course, constrained to be opposite).

The annular stimulus was used for various reasons. First, the motion stimulus was presented at a constant eccentricity in the periphery, and the effects of anisotropy of the retina were averaged across equivalent areas of the visual field. Second, it was easier to maintain fixation so eye movements were better controlled.* Finally, the use of circularly symmetric stimuli made it possible to produce a continuous animation by cyclically presenting a finite series of frames (computer memory efficiency).

---

*Torsional eye-movements induced by the rotating annuli (cyclo-induction) were not controlled in our experiment. Balliet and Nakayama (1978) reported the ability of extremely trained subjects to make stepwise eye torsions up to rotations of approx. 26 deg for large field stimuli (25–50 deg of visual angle). However, we do not expect torsional pursuit in our experimental conditions: small field stimuli, brief presentations, fast motion, unpredictable motion direction, and ambiguous or near-threshold motion stimuli.

Motion Competition Scheme II. Scheme II is very similar to Scheme I, except that textures s and v are interchanged. The motion stimulus and resulting motion paths for this experiment are sketched in Fig. 3

Although the heterogeneous motion path (between patches of texture s and v) is identical to that of Scheme I, the homogeneous motion path is different from that of Scheme I. In Scheme II, the homogeneous motion path consists of patches of texture v.

Textures

To study the effects that textural properties have on motion perception, we used a family of textures consisting of sinusoidally modulated grating patterns. The parameters that characterize these gratings were (1) modulation amplitude c, (2) spatial frequency ω, and (3) the orientation of the grating. The phase γ of the grating was a random variable uniformly distributed on the interval $[0, 2\pi]$.

The modulation was either tangential to the annulus or radial yielding two different orientations of the grating. Tangentially modulated "pinwheel" gratings are conveniently described by the luminance profile L_t in polar coordinates relative to the center of the annulus:

$$L_t(\phi, \rho) = L_0[1 + c \sin(2\pi\phi\omega\rho_0 + \gamma)]. \quad (1)$$

In this and following expressions, ϕ is the polar angle of a point in the image, ρ the distance to the center of the annulus, ω the spatial frequency (in c/deg) and L_0 the background luminance. Here ρ_0 is the distance (in deg of visual angle) between the center of the annulus and the center of the stimulus patches, and the pinwheel L_t runs through $2\pi\omega\rho_0$ cycles per circumference. This yields the desired effect that along the circumference that runs through the center of the stimulus annulus, the frequency of tangential modulation is ω c/deg.

Radially modulated "bull's eye" gratings are described by the luminance profile L_r:

$$L_r(\phi, \rho) = L_0[1 + c \sin(2\pi\omega \arctan(\rho/d) + \gamma)], \quad (2)$$

where d is the viewing distance.

The value of spatial frequency ω was varied over a range of three octaves: ω was 1.2, 4.9 or 9.9 c/deg. The amplitude c of the luminance modulation was adjustable between $c = 0$ (uniform patches) and $c = 1$ (maximum amplitude) in steps of 39 milli-units. At maximum amplitude, the minimum luminance value of gratings was 0 and the maximum luminance value of the gratings ($2L_0$) was equal to the maximum luminance value determined by the monitor setting. The phase (γ) of the modulation was randomized to insure that motion mechanisms sensitive to correspondence in stimulus luminance were not systematically engaged (Chubb & Sperling, 1988, 1991).

The orientation α of a grating is specified *relative* to the *tangent* to the annulus at the center of the grating. Tangentially modulated gratings L_t are called *orthogonal* to the their tangent and are labeled with orientation $\alpha = 90$ deg. Radially modulated gratings L_r are called *parallel* to their tangent and are labeled with orientation $\alpha = 0$ deg. The *direction* of motion between two successive gratings in a motion path is defined as the direction of the displacement of the centers of the gratings. It should be noted here, that the motion *direction* assigned to a grating deviates slightly (11.25 deg) from the tangent to the annulus at the center of this grating. In the following sections, however, we discuss gratings with $\alpha = 0$ deg as being oriented in the motion direction and gratings with $\alpha = 90$ deg as being oriented orthogonal to the motion direction.

An example of a series of frames for Scheme I, containing the textures described above, is shown in Fig. 4 in which texture s is a "medium" frequency
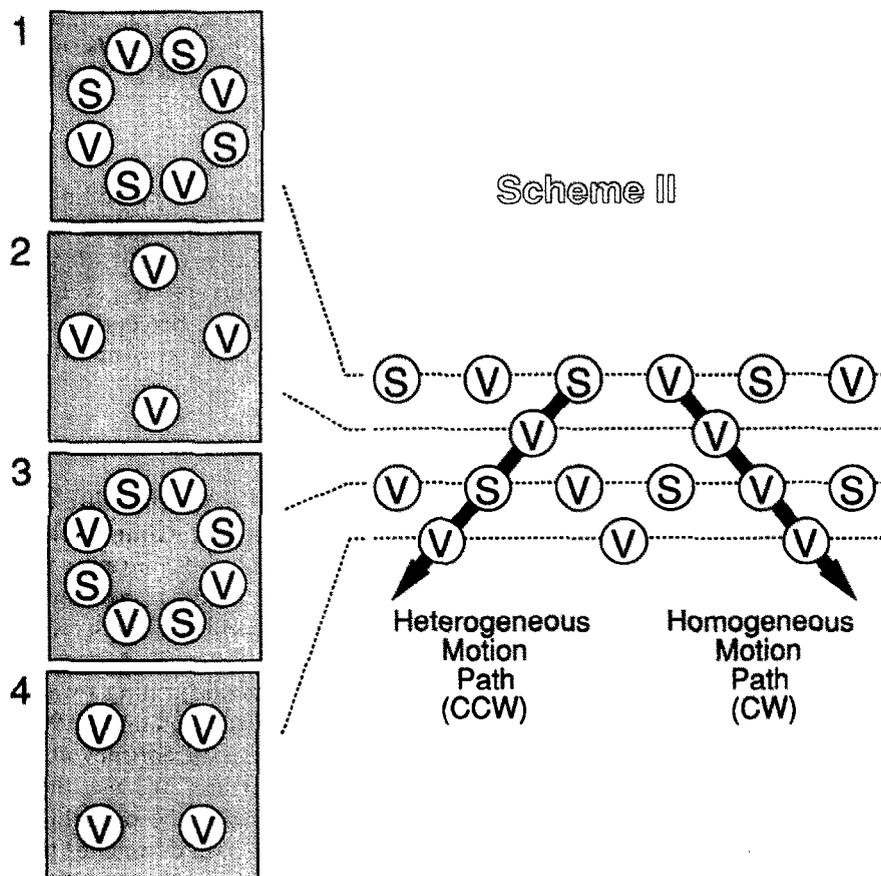

FIGURE 3. Motion competition Scheme II. This scheme is similar to Scheme I (see Fig. 2), except that textures s and v are interchanged. As a result the homogeneous motion path contains textures v.
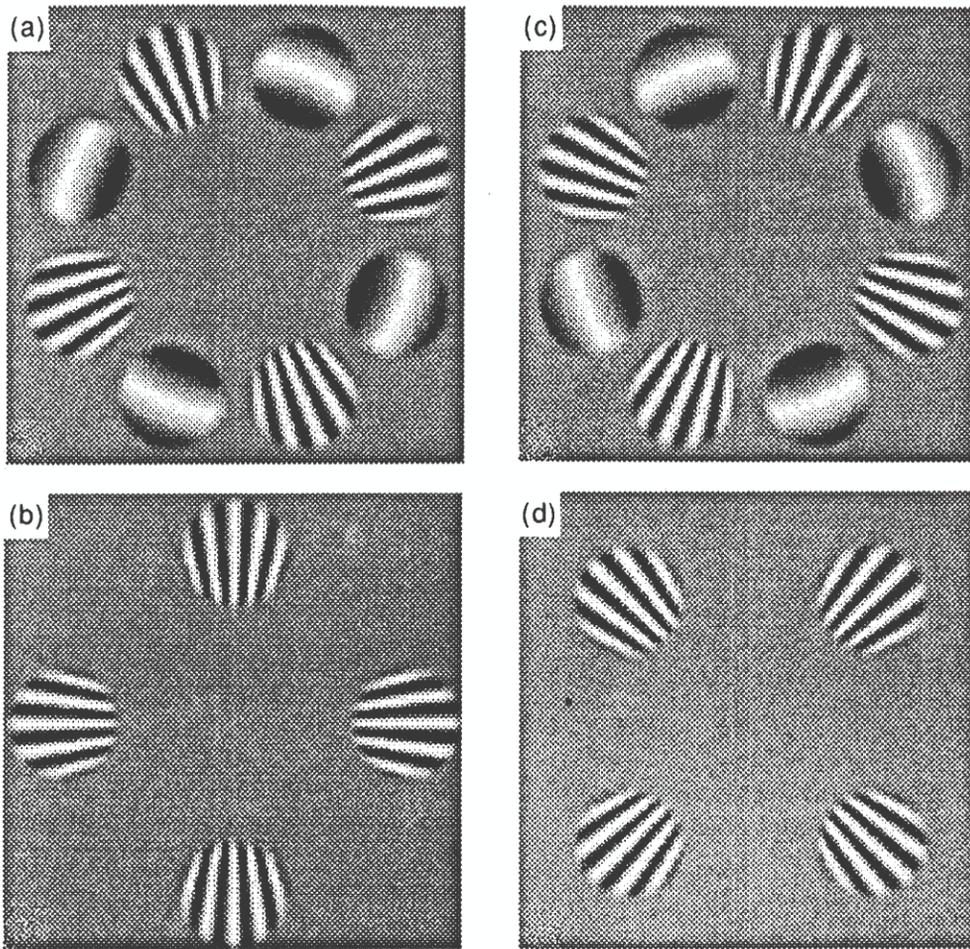
FIGURE 4. An example of the ambiguous motion display (as sketched in Fig. 2). Frames f_1, f_2, f_3, and f_4 (containing the patches of textures) are shown in (a), (b), (c) and (d) respectively. For this example, textures s and v differ two octaves in spatial frequency ($\omega_s = 4\omega_v$). Furthermore, the orientation of texture v is orthogonal to that of s.

orthogonal grating L_t (the standard texture) and in which texture v is a "low" frequency parallel grating L_r. The region outside the annulus (background) was uniform gray and had a luminance value ($L_0 = 72$ cd/m²). This background luminance was equal to the average luminance values across the gratings within the annulus.

Apparatus

The experiment was controlled by a IBM 386 PC compatible computer, driving a True Vision AT-Vista video graphics adapter. A 60 Hz Imtec 1261L monitor with a P4-type phosphor was used to display the stimuli. The screen dimensions were 21.8×14 cm (640×480 pixels; 12.3×8.0 deg).* The decay time to 10% and 1% intensity was about 1.3 and 6.2 msec respectively which is shorter than the temporal properties of retinal processing (Farrell *et al.*, 1990).

---

*Due to the limited bandwidth of the video amplifier (30 MHz) of the monitor, an anisotropy was observed for the average luminance of differently oriented textures that contain high spatial frequencies. Therefore, we only displayed the pixels at column position m and row position n for which ($m + n$) was even. The other pixels were dark. Hence, vertical and horizontal gratings shared a common "carrier" component. This procedure forfeits maximum luminance and resolution in favor of eliminating anisotropy; the net resolution (320×240 pixels) was more than adequate for the displays.

Subjects

Two subjects participated in the experiments: one of the authors (PW) and a colleague (JS). PW is emmetropic. JS is myopic (-0.5 D). Both subjects are experienced observers in psychophysical experiments related to motion perception. Natural pupils and spectacle corrections were used throughout.

Procedure

The subjects were asked to indicate the perceived direction of motion (counter-clockwise/clockwise) by pressing one of two buttons. All observations were binocular. The stimulus was viewed with visual fixation on a spot in the center of the screen at a viewing distance $d = 100$ cm.

In all experiments, grating s was a tangentially modulated grating (orthogonal, $\alpha_s = 90$ deg) with amplitude $m_s = 0.5$ and spatial frequency $\omega_s = 4.9$ c/deg. We call s the *standard* texture. We use the method of constant stimuli to determine the probability P that the heterogeneous motion path dominates the homogeneous path, as a function of the properties of the other grating v. The properties of grating v that were varied were orientation, α_v, amplitude m_v and spatial frequency ω_v. The other

stimulus parameter that was varied was frame duration τ.

$P_1(m_v; \omega_v, \alpha_v, \tau)$ is the probability of heterogeneous path dominance when using competition Scheme I as a function of amplitude m_v for a given spatial frequency ω_v and orientation α_v of v and a given frame duration τ. Similarly, we measured the probability of heterogeneous path dominance $P_2(m_v; \omega_v, \alpha_v, \tau)$ when using competition Scheme II. We examined three different spatial frequencies, $\omega_v = 1.2$, 4.9 and 9.9 c/deg; two orientations of gratings v, $\alpha_v = 0$ and 90 deg (parallel or orthogonal to the motion direction); and three frame durations, $\tau = 67$, 133 and 267 msec, corresponding to stimulus temporal frequencies 3.75, 1.88 and 0.94 Hz.

The probabilities $P_1(m_v; \omega_v, \alpha_v, \tau)$ and $P_2(m_v; \omega_v, \alpha_v, \tau)$ are measured as the percentage of heterogeneous motion percepts out of 36 presentations.

For a given spatial frequency, ω_v, and orientation, α_v, of grating v, and a given frame duration τ, we measured $P_1(m_v; \omega_v, \alpha_v, \tau)$ and $P_2(m_v; \omega_v, \alpha_v, \tau)$ as a function of amplitude m_v for both competition schemes in one session. Trials for these different conditions were mixed randomly in one session. Different spatial frequencies ω_v and different orientations α_v, were tested in different sessions.

MOTION COMPUTATIONS AND THEIR PREDICTIONS

The two competition schemes presented above comprise a powerful tool to discriminate between *single*-channel and *multi*-channel motion computations. Both computations, when applied to the competition schemes above, yield strong predictions for the results of the experiments presented in this paper. A key notion is transition invariance.

Single-channel motion computation: theory

In a previous paper (Werkhoven *et al.*, 1993), we showed the sufficiency of a single-channel motion computation for a two-parameter family of textures characterized by amplitude and spatial frequency. For the particular details of this computation (e.g. filter characteristics), see Werkhoven *et al.* (1993). Here we are concerned with the basic assumptions of a single-channel motion computation and their predictions for transitions using the competition schemes described above.

A single-channel motion computation model is based on the following assumptions.

● *Assumption 1.1*: the spatiotemporal input is pre-processed by a single nonlinear transformation T whose output we call *activity*.

We write T_s and T_v for the activities of textures s and v respectively. These activities can be written as functions of spatiotemporal properties of textures such as amplitude m, spatial frequency ω, orientation α and display duration τ: $T_s = T(m_s, \omega_s, \alpha_s, \tau)$ and $T_v = T(m_v, \omega_v, \alpha_v, \tau)$. Different textures can have identical activities; in other words they can correspond to

points of one isoactivity curve in the parameter space isomorphic to the family of textures examined.

● *Assumption 1.2*: the activity transformation T is then input for standard motion analysis.

The term "standard motion analysis" refers to the class of computations equivalent to some variant of the bilocal correlation scheme originally proposed by Reichardt and elaborated by van Santen and Sperling (1984). The computations in this class generally presume that the motion carried by a spatiotemporal signal is indicated by its Fourier energy spectrum. The Fourier components (drifting sinusoidal gratings) with the most energy dominate the motion elicited by that signal. For the highly restricted competition schemes we use in this paper, it is reasonable to assume that the strength of a motion path yielded by standard motion analysis is given by the *product* of the *activities* of two consecutive patches of texture in a motion path.

Single-channel motion computation: prediction

The motion strength of the heterogeneous path $S_{1,he}$ in Scheme I (between texture s and v) is the product of the activity T_s of texture s and the activity T_v of texture v:

$$S_{1,he} = T_s T_v. \tag{3}$$

The strength of the homogeneous path $S_{1,ho}$ in Scheme I (between identical textures s) is the square of the activity T_s of texture s. Heterogeneous motion was arbitrarily considered to have positive sign. Homogeneous motion ($S_{1,ho}$) has opposite polarity, since the motion is in the opposite direction:

$$S_{1,ho} = -T_s^2. \tag{4}$$

The strength of motion perception in the direction of the heterogeneous path is proportional to the sum of the output of standard motion analysis for the heterogeneous path and the homogeneous path.

Of particular interest is the condition in which the strength of the heterogeneous motion path ($S_{1,he}$) and the strength of the homogeneous motion paths ($S_{1,ho}$) are equal (except for opposite sign), i.e. when the two motion paths are balanced. Motion paths are balanced when the activities of texture s and v are equal: $T_s = T_v$. Hence, s and v correspond to points on one isoactivity curve in the parameter space.

The similarity of textures s and v is not what determines motion strength. When the motion paths are balanced, we can interchange textures s and v while the motion paths remain balanced in strength. This means that we will find that Schemes I and II will have identical activities of texture v at the point where the motion paths are balanced. We will refer to this property as *transition invariance*. Transition invariance is an exclusive property of a single channel motion computation, as we will see next.

Multi-channel motion computation: theory

A multi-channel motion computation differs from a single-channel motion computation in that standard

motion analysis is performed on more than one activity representation of the optical input.

● *Assumption* 2.1: the optical input is preprocessed by a set of n distinct activity transformations, T_i $(i = 1 \ldots n)$, yielding n different *activity representations* of the optical input.
● *Assumption* 2.2: standard motion analysis is applied to each of the activity representations.

An activity transformation followed by standard motion analysis is called a motion channel. Thus we have n different motion channels.

● *Assumption* 2.3: the outputs of the n channels are linearly summed with equal weights.

The net motion strength of a stimulus is proportional to the sum* of the outputs of standard motion analysis on each of the activity representations.

Multi-channel motion computation: predictions

An important consequence of multi-channel motion computations is that transition invariance is impossible for $n > 1$.

We first apply this motion computation to Scheme I (Fig. 2). For $i = 1, 2, \ldots, n$, we write $T_{is} = T_i(m_s, \omega_s, \alpha_s, \tau)$ and $T_{iv} = T_i(m_v, \omega_v, \alpha_v, \tau)$ for the outputs of transformation T_i applied to textures s and v respectively. The output of channel i for the heterogeneous path (between textures s and v) is the product of T_{is} and T_{iv}. The output of channel i for the homogeneous path (between textures s) is T_{is}^2.

We now derive an expression for the final motion strength after summing the outputs of different channels. For simplicity, we will use the vector notation:

$$
\mathbf{T}_s = \begin{bmatrix} T_{1s} \\ T_{2s} \\ \cdot \\ \cdot \\ T_{ns} \end{bmatrix} \quad \text{and} \quad \mathbf{T}_v = \begin{bmatrix} T_{1v} \\ T_{2v} \\ \cdot \\ \cdot \\ T_{nv} \end{bmatrix}. \tag{5}
$$

The vectors $\mathbf{T}_s$ and $\mathbf{T}_v$ are the activity vectors of textures s and v respectively. An activity vector represents the activity of a texture in the n-dimensional transformation space (T-space) defined by transformations $T_1 \ldots T_n$. For Scheme I the summed motion strengths of all channels for the heterogeneous path can be written as the scalar product:

$$
S_{1,\text{he}} = \mathbf{T}_s \cdot \mathbf{T}_v = \sum_{i=1}^{n} T_{is} T_{iv}. \tag{6}
$$

The motion strength of the homogeneous path is:

$$
S_{1,\text{ho}} = -\mathbf{T}_s \cdot \mathbf{T}_s = -\sum_{i=1}^{n} T_{is} T_{is}. \tag{7}
$$

---

*For specificity, but without loss of generality, we have assumed summation with equal weights. In fact, we can allow different weights when they are all *positive* (different positive weights can be absorbed in the different transformations T_i).

As a result, the net motion strength (D_1) in the direction of the heterogeneous path is:

$$
D_1 = S_{1,\text{he}} + S_{1,\text{ho}} = \mathbf{T}_s \cdot [\mathbf{T}_v - \mathbf{T}_s]. \tag{8}
$$

At a transition, the motion strength of the heterogeneous and homogeneous motion paths are balanced, i.e. the net motion strength is zero:

$$
D_1 = \mathbf{T}_s \cdot [\mathbf{T}_v - \mathbf{T}_s] = 0. \tag{9}
$$

In T-space, there exists an $(n - 1)$-dimensional solution for $\mathbf{T}_v$ in equation (9). For example, we find a point when $n = 1$, a straight line (perpendicular to $\mathbf{T}_s$) when $n = 2$, and a plane (orthogonal to $\mathbf{T}_s$) when $n = 3$.

It should be noted that the net heterogeneous motion strength D_1 can be positive. Hence, even in a multi-channel motion computation, the strength of the heterogeneous motion path can dominate.

Using only the result for competition Scheme I, we cannot discriminate between a *single-channel* and *multi-channel* motion computation, because both computations produce transitions from dominance of the homogeneous motion path to dominance of the heterogeneous motion path as the projection of $\mathbf{T}_v$ on $\mathbf{T}_s$ increases. We will show that *transition invariance* is critical in discriminating a *single-channel* from a *multi-channel* motion computation.

Using analogous notation and reasoning, we find the solution for vectors $\mathbf{T}_v$ for which the motion strength of the heterogeneous and homogeneous paths are balanced when using competition Scheme II (Fig. 3):

$$
D_2 = \mathbf{T}_v \cdot [\mathbf{T}_s - \mathbf{T}_v] = 0. \tag{10}
$$

Note that Scheme II results from Scheme I by interchanging elements s and v. In T-space there exists a $(n - 1)$-dimensional solution for $\mathbf{T}_v$ in equation (10). For example, we find a point when $n = 1$ $(\mathbf{T}_v = \mathbf{T}_s)$, a circle (tangent to the line perpendicular to $\mathbf{T}_s$) when $n = 2$, and a spherical surface (tangent to the plane that is orthogonal to $\mathbf{T}_s$) when $n = 3$.

Transition invariance

The transition points for Scheme I will be invariant after interchanging elements s and v (resulting in Scheme II) when the activity vector $\mathbf{T}_v$ satisfies both equations (9) and (10). As is easily seen, there is only a single activity vector $\mathbf{T}_v$ that satisfies both equations: $\mathbf{T}_v = \mathbf{T}_s$.

Vector $\mathbf{T}_v$ is equal to vector $\mathbf{T}_s$ when each transformation T_i involved in the motion computation has an equal output for both textures v and s:

$$
T_{is} = T_{iv} \ (i = 1 \ldots n). \tag{11}
$$

This is a strong constraint for the ensemble of transformations that might be involved in a *multi-channel* motion computation. Every transformation T_i must yield equal activity for both textures s and v, i.e. both textures map to an isoactivity contour of T_i. Furthermore, when transition invariance holds for a range of properties of texture v, the isoactivity contours of each transformation T_i must be identical within this range.

This means that the motion computation is one-dimensional, i.e. we have a *single-channel* motion computation.

In conclusion, transition invariance for a range of texture properties excludes a motion computation involving more than one transformation of the optical input (or more than one texture grabber). Transition invariance can be tested by comparing the conditions for which the motion paths are balanced between Schemes I and II.

EXPERIMENTS

We measured the probability $P_1(m_v; \omega_v, \alpha_v, \tau)$ of heterogeneous path dominance using Scheme I, and the probability $P_2(m_v; \omega_v, \alpha_v, \tau)$ using Scheme II. The probabilities P_1 and P_2 are measured as a function of amplitude m_v of grating v for 18 conditions: three different spatial frequencies of grating v, two orientations of grating v and three frame durations. For each scheme we will show the probability curves for each of these conditions.

The effects of spatial frequency on the strength of apparent motion perception have been presented and discussed in detail in another paper (Werkhoven *et al.*, 1993). In this paper, the organization of the presentation is chosen to facilitate the analysis of the effects of orientation and frame duration. Hence, for one particular spatial frequency ω_v, we grouped together the probability curves $P_i(m_v; \omega_v, \alpha_v, \tau)$, for $i = 1,2$ (both competition schemes), $\alpha_v = 0, 90$ deg (both orientations), and $\tau = 67, 133$ or 267 msec (all three frame durations) in the same figure (see e.g. Fig. 5).

For each of the probability curves $P_i(m_v; \omega_v, \alpha_v, \tau)$, we write $\mu_i(\omega_v, \alpha_v, \tau)$ for the amplitude m_v satisfying $P_i(m_v; \omega_v, \alpha_v, \tau) = 50\%$. Thus, $\mu_i(\omega_v, \alpha_v, \tau)$ is the amplitude for which the strength of the heterogeneous motion path and the strength of the homogeneous motion path are balanced. The amplitudes μ_1 and μ_2 are called the *transition amplitudes* for Schemes I and II respectively. Although μ_1 and μ_2 are both functions of ω_v (spatial frequency of v), α_v (orientation of v), and τ (frame duration of the presentation), it will be convenient to omit explicit reference to all but the argument(s) relevant
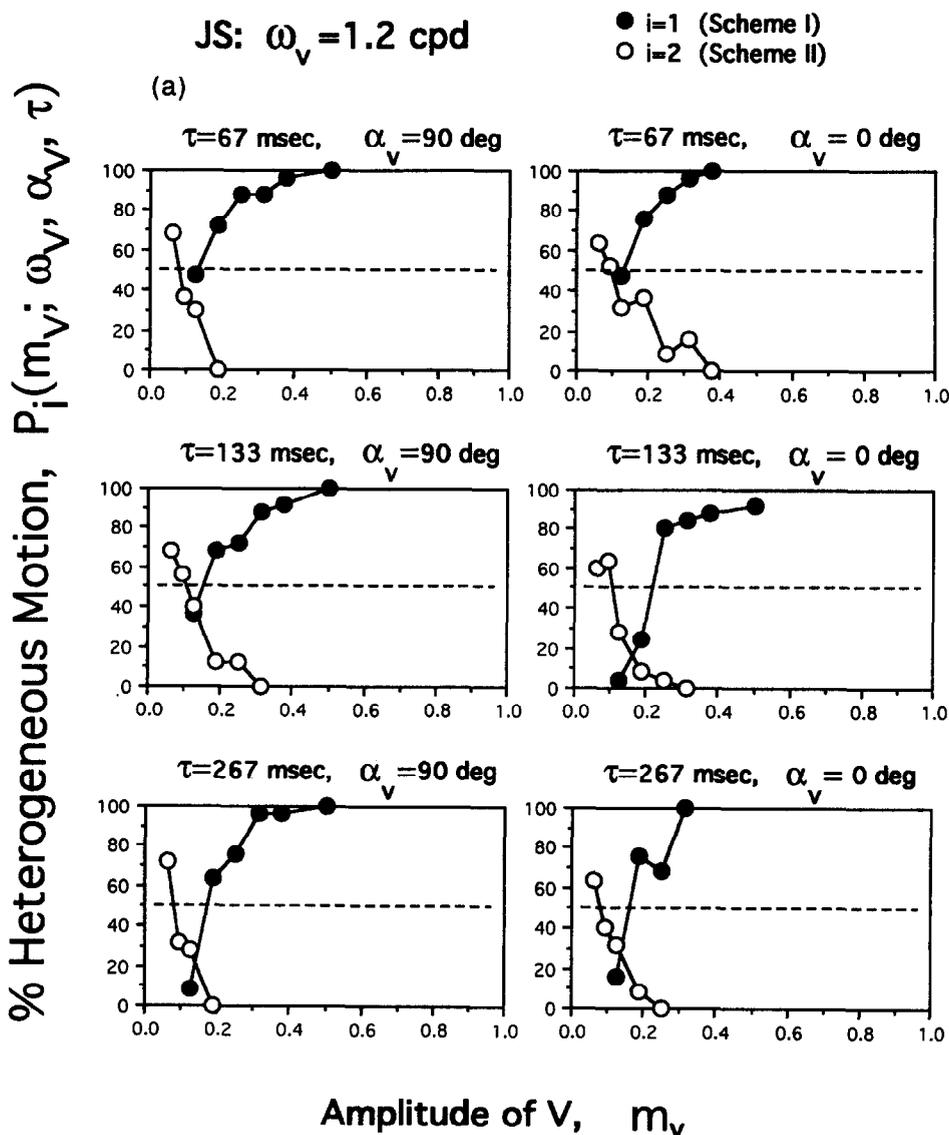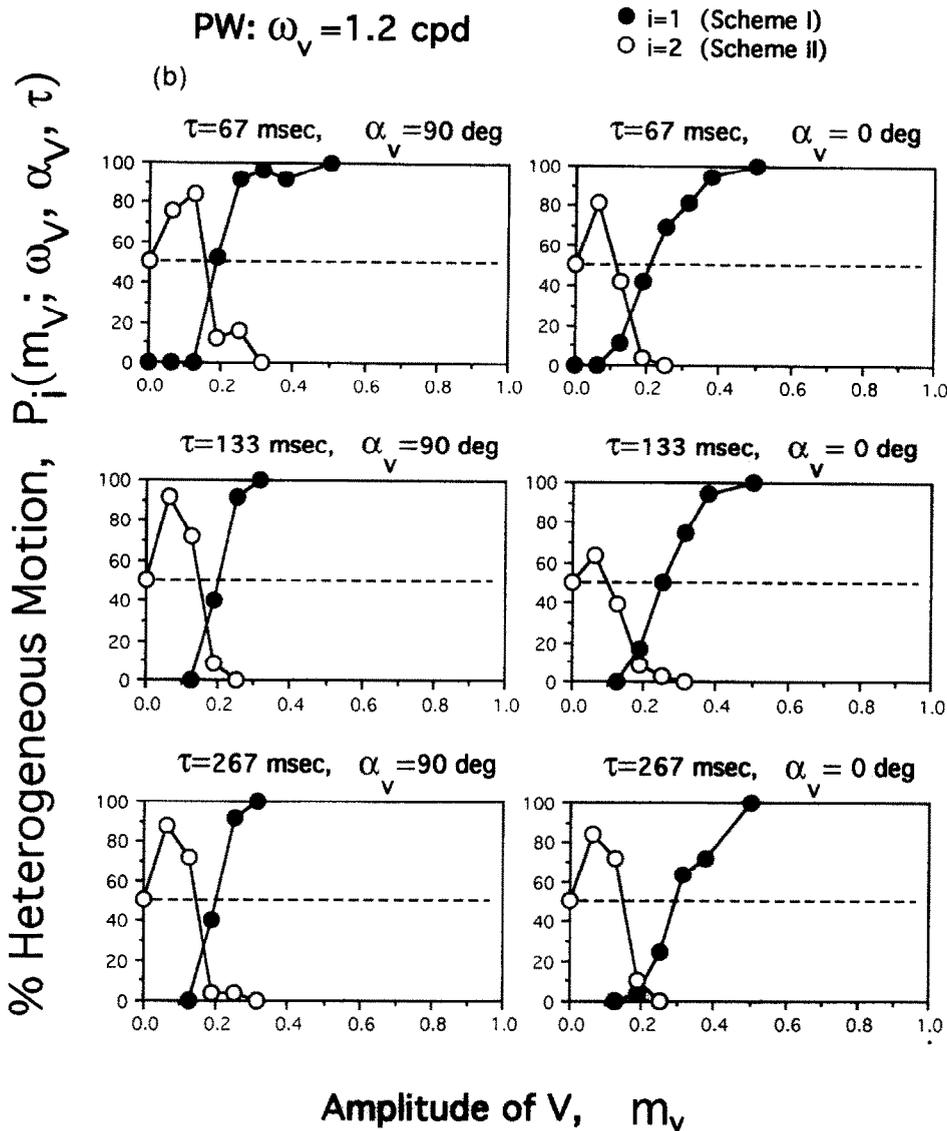

FIGURE 5(a). *Caption overleaf.*

FIGURE 5. Effect of frequency and orientation differences: the probability $P_i(m_v; \omega_v = 1.2\,c/deg, \alpha_v, \tau)$ of heterogeneous path dominance as a function of the amplitude m_v of texture v for two different orientations ($\alpha_v = 0\,deg$ and $\alpha_v = 90\,deg$) of texture v and for different frame durations ($\tau = 67$, 133 and 267 msec) and two subjects. Solid circles represent the probability P_1 for Scheme I (Fig. 2), open circles P_2 for Scheme II (Fig. 3). The horizontal dashed line indicates a 50% probability of heterogeneous path dominance. The amplitude m_s, spatial frequency ω_s, and orientation α_s of texture s is the same for all panels: $m_s = 0.5$, $\omega_s = 4.9\,c/deg$, and $\alpha_s = 90\,deg$.

in a given context whenever the omission causes no confusion.

Estimation of transition amplitudes μ_1 and μ_2

To estimate transition amplitude $\mu_i(\omega_v, \alpha_v, \tau)$, we selected the three data points of each probability curve around the crossing of the curves with the 50% guide line that were closest to the 50% guide line. Within this selected range, the curve was assumed to be linear, and these data points were subject to a least-squares method of linear regression.

A Monte Carlo bootstrapping method was used to estimate the variability in the measurements of $\mu = \mu_1, \mu_2$. Specifically, let p_1, p_2 and p_3 be the estimates, obtained from the data, of the probability of heterogeneous motion for the three amplitudes x_1, x_2 and x_3 in

the neighborhood of the transition amplitude, μ. Each estimated p_i is derived from 36 trials at amplitude x_i. We assume that p_1, p_2 and p_3 are the p-parameters of independent, 36-trial, binomial random variables underlying our data. Then we (a) obtain 400 realizations of each of three binomial random variables, $n_1(N = 36, p = p_1)$, $n_2(N = 36, p = p_2)$, and $n_3(N = 36, p = p_3)$; (b) fit a straight line to the obtained points $(x_1, n_1/N)$, $(x_2, n_2/N)$, and $(x_3, n_3/N)$; and (c) find the amplitude μ_{sim} for which (μ_{sim}, 0.5) is on this line. This procedure yields 400 independent, simulated estimates of the transition amplitude μ. The standard deviation of this distribution is taken as the standard error of our original estimate of μ.

The error in the estimation of $\mu_i(\omega_v, \alpha_v, \tau)$ depends on the steepness of the probability curve at the transition

TABLE 1. Transition amplitudes for observer PW

	$\omega_v = 1.2$ c/deg				$\omega_v = 4.9$ c/deg				$\omega_v = 9.9$ c/deg			
	$\alpha_v = 90$ deg		$\alpha_v = 0$ deg		$\alpha_v = 90$ deg		$\alpha_v = 0$ deg		$\alpha_v = 90$ deg		$\alpha_v = 0$ deg	
	μ_1	μ_2	μ_1	μ_2	μ_1	μ_2	μ_1	μ_2	μ_1	μ_2	μ_1	μ_2
$\tau = 67$ msec	0.19	0.16	0.21	0.11	0.52	0.52	0.62	0.49	0.75	0.68	0.92	0.78
$\tau = 133$ msec	0.20	0.15	0.26	0.10	0.49	0.51	0.55	0.42	0.68	0.62	0.89	0.79
$\tau = 267$ msec	0.20	0.15	0.31	0.15	0.50	0.50	0.50	0.42	0.59	0.58	0.92	0.83

Transition amplitude μ_1 for Scheme I and μ_2 for Scheme II are given for 18 conditions: three frame durations (horizontal rows), three spatial frequencies and two orientations (columns).

amplitude. As will be clear from the following Results section, the steepness of the probability curves is roughly determined by the spatial frequency ω_v. Therefore, we can summarize the errors in μ by taking the average errors which are 0.014 for the $\omega_v = 1.2$ c/deg condition, 0.022 for the $\omega_v = 4.9$ c/deg condition and 0.030 for the $\omega_v = 4.9$ c/deg condition.*

We will say that transition amplitudes μ_1 and μ_2 are significantly different (with a probability of 95%) when they differ more than 2 times the error in the difference of μ_1 and μ_2. That is a difference of more than 0.039 for the $\omega_v = 1.2$ c/deg condition, 0.062 for the $\omega_v = 4.9$ c/deg condition and 0.084 for the $\omega_v = 4.9$ c/deg condition.

Spatial frequency $\omega_v = 1.2$ c/deg

In Fig. 5, we show the probability curves for Scheme I (solid circles) and Scheme II (open circles) for which the spatial frequency of grating v is 2 octaves lower than

*On request of an anonymous reviewer, we have also applied a maximum likelihood estimation technique (called ERF-fit) using a cumulative normal distribution to model the probability curves (see Werkhoven et al., 1993). Cumulative normal distributions $Erf[(m_v - \mu_1)/\sigma_1]$ are theoretically justified for the probability curves obtained for Scheme I, but not for Scheme II (see Werkhoven et al., 1993). The value σ characterizes the steepness of the probability curve at the transition amplitude. No linearity assumption is required and all data-points can be used to estimate transition amplitude μ_1. These estimates obtained by ERF-fit are equal to those obtained by a linear fit within 1 SD for the ERF-fit estimates. The standard deviations for the ERF-fit estimates also depend on the steepness σ_1 of the probability curve and thus on the spatial frequency ω_v. Bootstrap simulations similar to those described above (but now taking all data points) yield standard deviations for the ERF-fit estimates of μ approx. 0.009, 0.012 and 0.019 for the $\omega_v = 1.2$, 4.9 and 9.9 c/deg conditions respectively. Because of the good agreement between the ERF-fit estimates and those of a linear fit and because the ERF-fit procedure cannot be applied directly to Scheme II, we have chosen to use the estimates of the linear fit procedure.

that of grating s ($\omega_v = 1.2$ c/deg). For each subject [JS, Fig. 5(a); PW, Fig. 5(b)] we collected data for 12 conditions. The 12 conditions shown result from two different orientations of grating v (left column, $\alpha_v = 90$ deg; right column, $\alpha_v = 0$ deg), and three different frame durations (top row, $\tau = 67$ msec; middle row, $\tau = 133$ msec; bottom row, $\tau = 267$ msec).

The curve marked by solid circles in the left upper panel of Fig. 5 shows that the probability $P_1(m_v;$ 1.2 c/deg, 90 deg, 67 msec) of heterogeneous path dominance increases monotonically with m_v. For sufficiently high amplitude ($m_v > 0.3$), heterogeneous motion dominates homogeneous motion in more than 90% of the presentations.

Following the estimation method, we find a transition amplitude $\mu_1(1.2$ c/deg, 90 deg, 67 msec) = 0.19 (subject PW) and 0.13 (subject JS) for Scheme I (see Tables 1 and 2).

Also in the upper left panel, we show the probability $P_2(m_v;$ 1.2 c/deg, 90 deg, 67 msec) of heterogeneous path dominance for Scheme II (open circles). When the amplitude of grating v is zero, P_2 is 50% (the competition scheme is purely ambiguous!). Starting at $m_v = 0$ and increasing m_v we find that P_2 first increases. It reaches some maximum and decreases monotonically when we further increase the amplitude m_v of grating v. We find a transition amplitude when P_2 is 50%. The transition amplitude for Scheme II is: $\mu_2(1.2$ c/deg, 90 deg, 67 msec) = 0.16 (subject PW) and 0.08 (subject JS). Thus, the transition amplitude of Scheme I is not markedly different from that of Scheme II (see Table 1).

Interestingly, the results are not much different for longer frame durations, as can be concluded from comparing the panels in the left column. The middle panel shows the probabilities $P_1(m_v;$ 1.2 c/deg, 90 deg, 133 msec) and $P_2(m_v;$ 1.2 c/deg, 90 deg, 133 msec) for a longer frame duration of 133 msec. For the left bottom

TABLE 2. Transition amplitudes for observer JS

	$\omega_v = 1.2$ c/deg				$\omega_v = 4.9$ c/deg				$\omega_v = 9.9$ c/deg			
	$\alpha_v = 90$ deg		$\alpha_v = 0$ deg		$\alpha_v = 90$ deg		$\alpha_v = 0$ deg		$\alpha_v = 90$ deg		$\alpha_v = 0$ deg	
	μ_1	μ_2	μ_1	μ_2	μ_1	μ_2	μ_1	μ_2	μ_1	μ_2	μ_1	μ_2
$\tau = 67$ msec	0.13	0.08	0.13	0.10	0.48	0.48	0.43	0.37	0.57	0.53	0.79	0.62
$\tau = 133$ msec	0.17	0.10	0.21	0.11	0.43	0.39	0.50	0.41	0.59	0.56	0.72	0.53
$\tau = 267$ msec	0.19	0.08	0.18	0.08	0.50	0.49	0.48	0.37	0.58	0.58	0.73	0.59

Transition amplitude μ_1 for Scheme I and μ_2 for Scheme II are given for 18 conditions: three frame durations (horizontal rows), three spatial frequencies and two orientations (columns).

panel the frame duration was 267 msec. The results are nearly identical for all three frame durations.

In the right column, we show the results for three different frame durations when grating v is oriented *parallel* to the motion direction (while grating s is still oriented orthogonal to the motion direction). Now, textures s and v of the heterogeneous motion path have perpendicular orientations. Although gratings s and v differ two octaves in spatial frequency and have perpendicular orientations, the heterogeneous path can easily dominate the homogeneous path for sufficient amplitude m_v. However, the right upper panel shows that, for a brief frame duration, the transitions $\mu_1(1.2 \text{ c/deg}, 0 \text{ deg}, 67 \text{ msec})$ and $\mu_2(1.2 \text{ c/deg}, 0 \text{ deg}, 67 \text{ msec})$ are slightly but significantly different from those found when gratings s and v have identical orientations (compare the right upper panel with the left upper panel).

Also at longer frame durations (middle right and bottom right panels), heterogeneous motion easily dominates over homogeneous motion for sufficiently high

amplitude m_v. Again we find that the transitions μ_1 and μ_2 are slightly but significantly shifted apart for different oriented gratings s and v.

Spatial frequency $\omega_v = 4.9 \, c/deg$

The effects of grating orientation on the strength of apparent motion is most clearly revealed by letting gratings s and v have identical spatial frequencies and by comparing a condition in which the orientations of both gratings are identical with a condition in which gratings s and v are perpendicular. Figure 6 shows the results when gratings s and v have an identical spatial frequency of 4.9 c/deg. Similar to Fig. 5, the different rows show results for different frame durations.

In the left column, both spatial frequency and orientation of grating s and v are identical. As expected, transitions occur when the amplitude m_v of grating v is equal to the amplitude of gratings s [μ_1 (4.9 c/deg, 90 deg, τ) = μ_2(4.9 c/deg, 90 deg, τ) = 0.5]. At this amplitude, gratings s and v are identical (except for random phase differences).
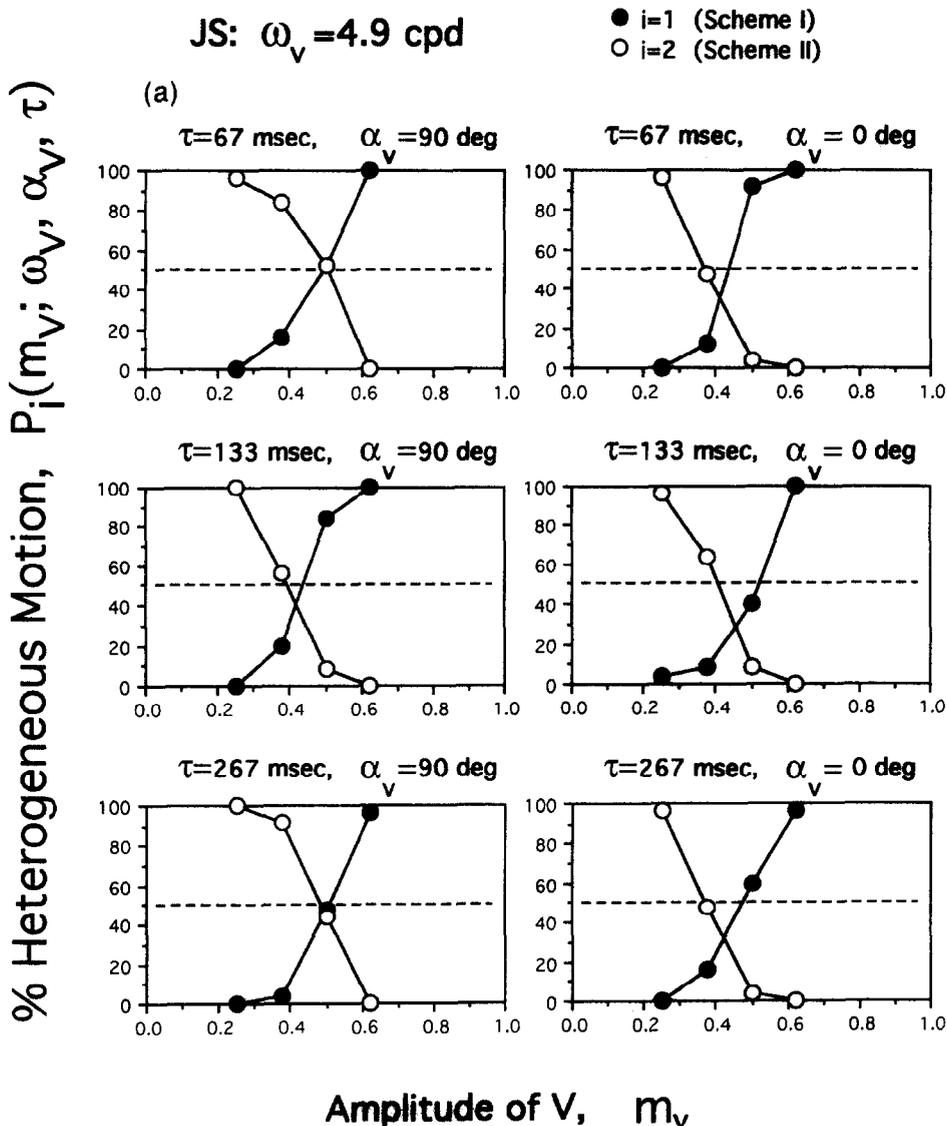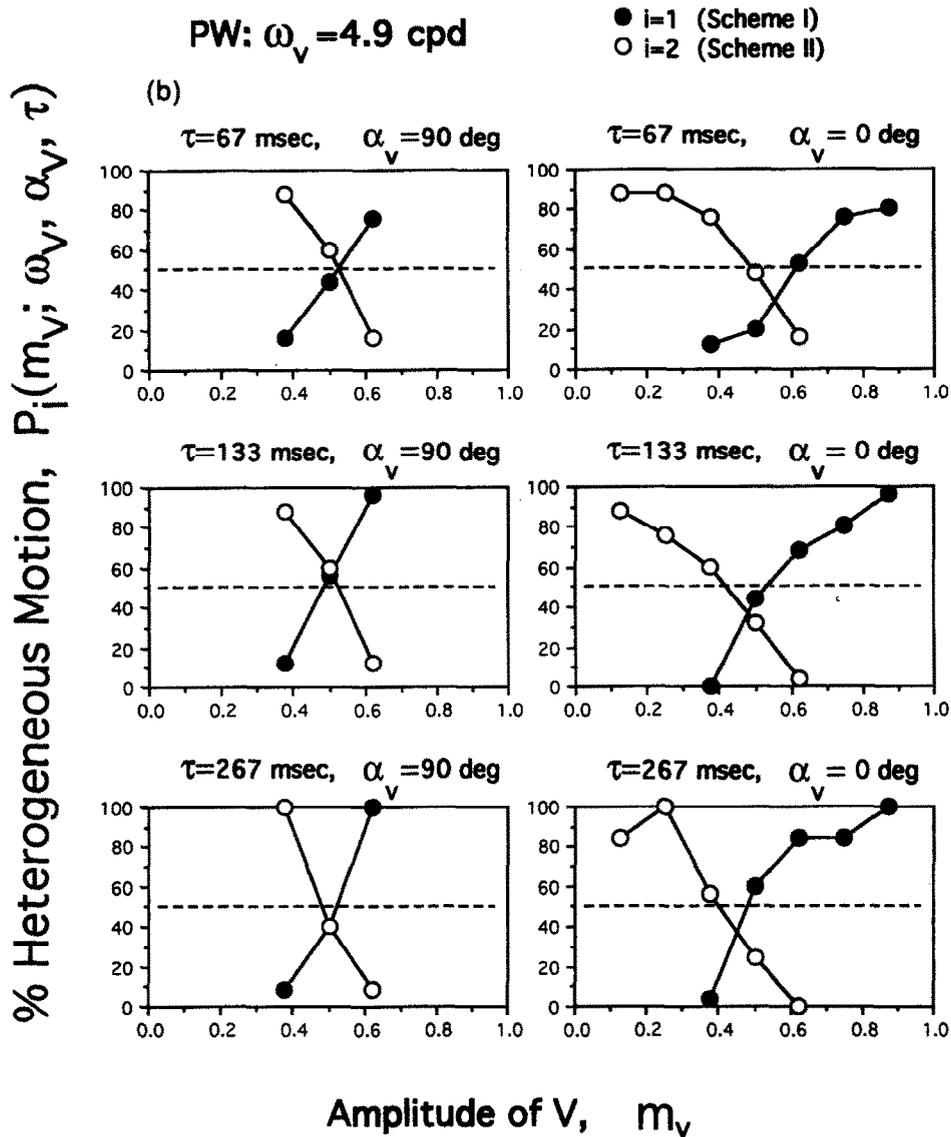

FIGURE 6(a). *Caption on facing page.*

FIGURE 6. Effect of orientation differences: the probability $P_i(m_v; \omega_v = 4.9\,c/\mathrm{deg}, \alpha_v, \tau)$ of heterogeneous path dominance as a function of the amplitude m_v of texture v for two different orientations ($\alpha_v = 0\,\mathrm{deg}$ and $\alpha_v = 90\,\mathrm{deg}$) of texture v and for different frame durations ($\tau = 67$, 133 and 267 msec) and two subjects. Solid circles represent the probability P_1 for Scheme I (Fig. 2), open circles P_2 for Scheme II (Fig. 3). The horizontal dashed line indicates a 50% probability of heterogeneous path dominance. The amplitude m_s, spatial frequency ω_s and orientation α_s of texture s is the same for all panels: $m_s = 0.5$, $\omega_s = 4.9\,c/\mathrm{deg}$, and $\alpha_s = 90\,\mathrm{deg}$.

By comparing the left and right panel in a row, we can see the effect of orientation differences between gratings s and v on the relative strength of the heterogeneous path. Obviously, the transition amplitudes μ_1 (4.9 c/deg, 0 deg, τ) and μ_2(4.9 c/deg, 0 deg, τ) are slightly but significantly shifted apart when the orientations of grating s and v are perpendicular (see right column). The transition amplitude for Scheme I is systematically higher than that for Scheme II. This effect of orientation is similar for different frame durations (see Table 1).

Spatial frequency $\omega_v = 9.9\,c/deg$

Figure 7 shows the results for a grating v with a spatial frequency of 9.9 c/deg (1 octave higher than the spatial frequency of grating s). For this condition, we need transition amplitudes m_v that are higher than that of

grating s to have heterogeneous path dominance. In the left column, the differences between transitions μ_1(9.9 c/deg, 90 deg, τ) and μ_2(9.9 c/deg, 90 deg, τ) are insignificant.

Looking at the right columns (perpendicularly oriented gratings s and v), we find that the transitions are slightly but significantly shifted apart.

MODEL IMPLICATIONS

We start out with a rough analysis of the results for Schemes I and II separately. Subsequently, we'll elaborate our analysis to account for the slight difference between the results found with Scheme I and those with Scheme II.

Rough analysis of Scheme I

The results of the experiments with Scheme I show that heterogeneous motion is dominant over homogeneous motion whenever the activity T_v of grating v exceeds some value (i.e. when m_v exceeds the transition μ_1). At the transitions, heterogeneous and homogeneous motion strength is balanced. These transitions are approximately independent of frame duration τ. This result for Scheme I is well explained in terms of a single-channel motion computation as described earlier. The motion strength $S_{1,he}$ of the heterogeneous path containing gratings s and v is proportional to the product of the outputs of a certain transformation T of these gratings: $S_{1,he} = T_s T_v$. The motion strength $S_{1,ho}$ of the homogeneous path is: $S_{1,ho} = T_s T_s$. After linear combination, the net motion strength in the direction of the heterogeneous motion path is: $D_1 = T_s (T_v - T_s)$ [see equation (10)]. Thus, the net motion strength is linear in the activity T_v of texture v. Furthermore, we have shown (see Werkhoven *et al.*, 1993) that the activity T is linear

in amplitude m: $T(m, \omega, \alpha, \tau) = mf(\omega, \alpha)$. Assuming a Normal noise distribution with mean D_1 for the measurement of net motion strength, our model predicts that the probability $P_1(m_v; \omega_v, \alpha_v, \tau)$ is an error-function of m_v. The *shapes* of the probability curves found for Scheme I (see e.g. Fig. 5) are qualitatively in agreement with this simple model.

Further, we consider the *transitions* of the probability curves. At a transition amplitude the motion strengths of both paths are balanced, i.e. the activities of both gratings are equal: $T_v(m_v, \omega_v, \alpha_v, \tau) = T_s(m_s, \omega_s, \alpha_s, \tau)$. Since activity T_s is constant throughout the experiment, the dependence of activity T_v on spatial frequency ω_v and amplitude m_v can be inferred from the transition amplitudes $\mu_i(\omega_v, \alpha_v, \tau)$ for different spatial frequencies but constant orientation $\alpha_v = a$ and constant frame duration $\tau = b$ (see Table 1). That is, we find: $f(\omega_v, a) = T_s \mu_i^{-1}(\omega_v, a, b)$. Obviously, transition amplitude $\mu_i(\omega_v, a, b)$ increases with frequency ω_v for all pairs a, b, suggesting that the function f and thus the activity T is a monotonically decreasing function of ω_v.
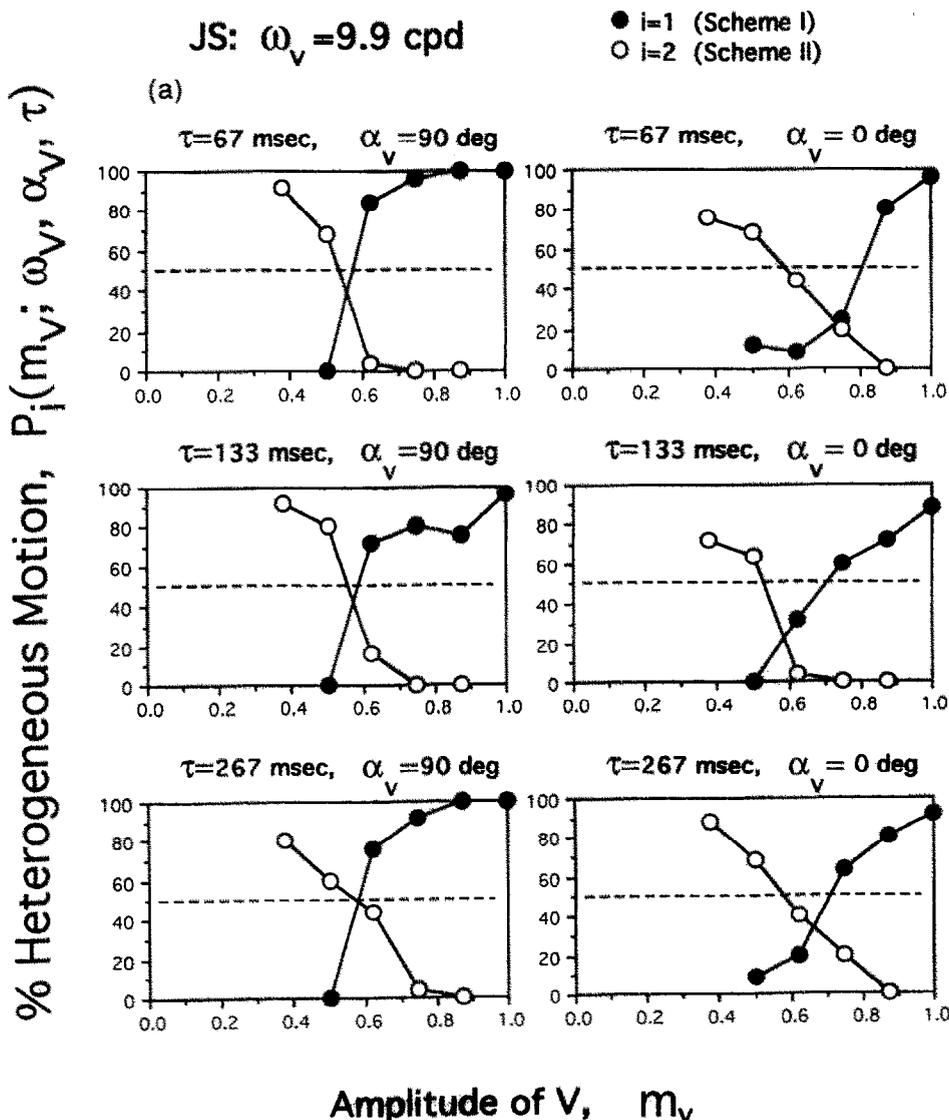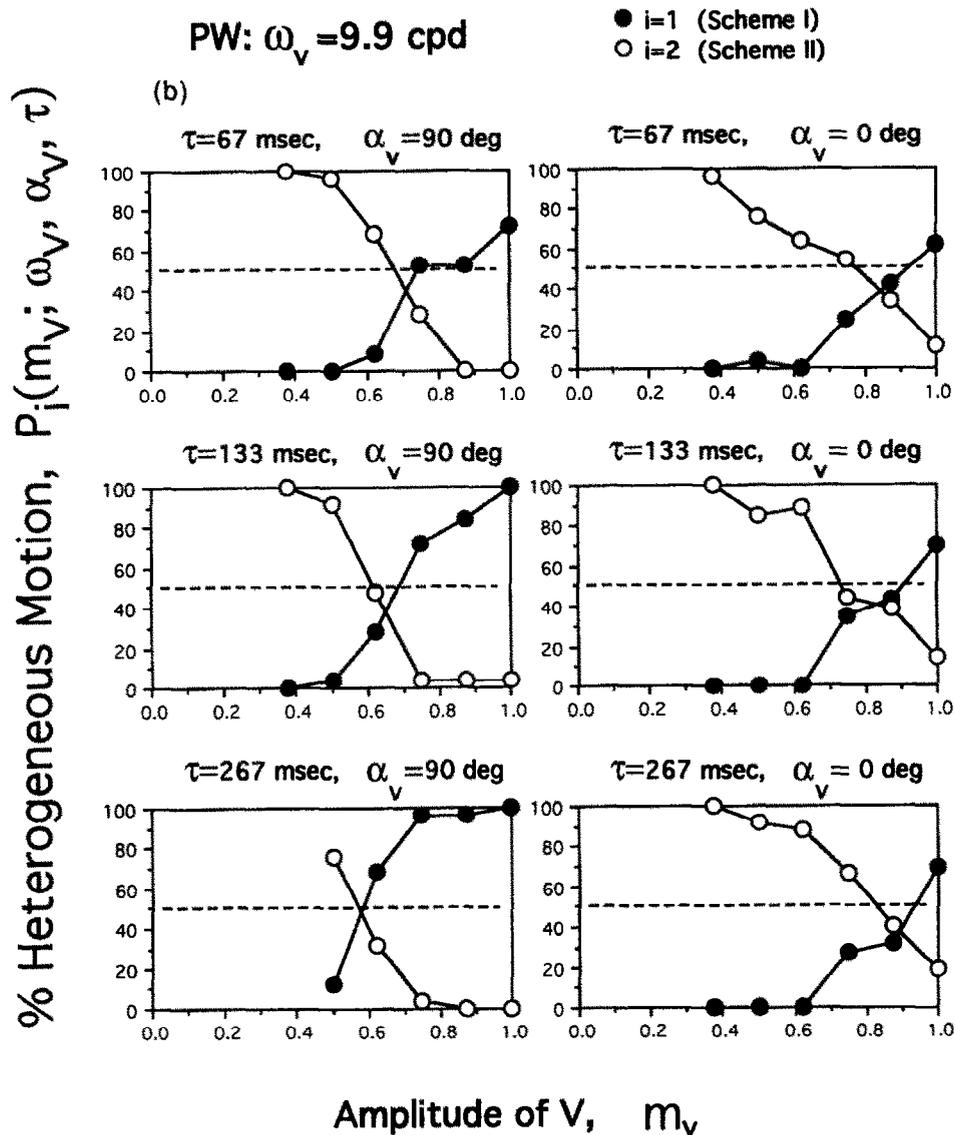

FIGURE 7(a). *Caption on facing page.*

PW: ω_v =9.9 cpd

● i=1 (Scheme I)
○ i=2 (Scheme II)

(b)

$P_i(m_v; \omega_v, \alpha_v, \tau)$ % Heterogeneous Motion,

τ=67 msec, α_v =90 deg

τ=67 msec, α_v = 0 deg

τ=133 msec, α_v =90 deg

τ=133 msec, α_v = 0 deg

τ=267 msec, α_v =90 deg

τ=267 msec, α_v = 0 deg

Amplitude of V, m_v

FIGURE 7. Effect of frequency and orientation differences: the probability $P_i(m_v; \omega_v = 9.9$ c/deg, $\alpha_v, \tau)$ of heterogeneous path dominance as a function of the amplitude m_v of texture v for two different orientations ($\alpha_v = 0$ deg and $\alpha_v = 90$ deg) of texture v and for different frame durations ($\tau = 67$, 133 and 267 msec) and two subjects. Solid circles represent the probability P_1 for Scheme I (Fig. 2), open circles, P_2 for Scheme II (Fig. 3). The horizontal dashed line indicates a 50% probability of heterogeneous path dominance. The amplitude m_s, spatial frequency ω_s, and orientation α_s of texture s is the same for all panels: $m_s = 0.5$, $\omega_s = 4.9$ c/deg, and $\alpha_s = 90$ deg.

Even when the gratings are orthogonally oriented, the strength of the heterogeneous motion path easily dominates the strength of the homogeneous path. The transition amplitudes for orthogonally oriented gratings v are approximately equal to those for parallel oriented gratings v for all spatial frequencies examined. This strongly suggests that the function f, and thus the activity T, is almost independent of its orientation α.

Finally, the results are similar for different frame durations, suggesting that f is independent of τ in the first approximation. In fact, the F-ratios given by an ANOVA test show that spatial frequency accounts for 87% of the variance in μ_1 for subject PW (95% for subject JS). Orientation accounts for only 11% (PW) and 3.5% (JS).

A tentative conclusion is that the results of Scheme I can be explained by standard motion analysis applied to

a single activity representation. The nonlinear transformation that yields the activity representation increases monotonically with amplitude and decreases monotonically with spatial frequency. Furthermore, the transformation is approximately rotationally invariant (not tuned to specific orientations of the texture sensed).

Rough analysis of Scheme II

If the conclusions for Scheme I were valid, the activity T_v of texture v would increase monotonically with its amplitude m_v. Applying our model to Scheme II, the strength of the homogeneous path increases quadratically with T_v, whereas the strength of the heterogeneous path increases linearly with T_v. The net motion strength in the direction of the heterogeneous path is $T_v(T_s - T_v)$ [see equation (10)]. As a result, the heterogeneous path
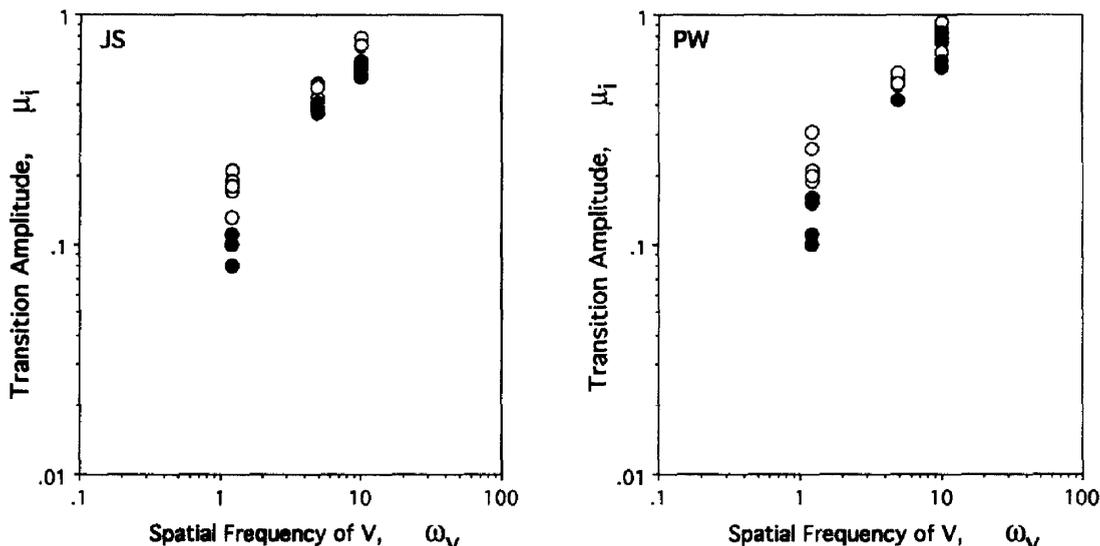
FIGURE 8. Transition amplitudes μ_i as a function of spatial frequency ω_v. In fact, we have projected the transition amplitudes which are represented in a four-dimensional graph—$\mu_i(\omega_v, \alpha_v, \tau)$—on the μ_i, ω_v-plane. Transitions for Scheme I ($i = 1$) are indicated by open circles, those for Scheme II ($i = 2$) by solid circles.

is dominant for small T_v and the homogeneous path is dominant for large amplitude T_v. As in Scheme I, the motion paths are balanced when $T_v = T_s$. This prediction is supported by the data. The curves start at $P_2 = 50\%$ climb to a maximum and decrease monotonically for larger amplitude m_v, crossing the dashed line ($P_2 = 50\%$) at the transition amplitude μ_2.

Consistent with the model and the results for Scheme I, we find that (1) the transitions μ_2 increase with increased spatial frequency; (2) the transitions μ_2 for orthogonally oriented gratings v are approximately equal to the transitions for parallel oriented gratings v for all spatial frequencies examined; and (3) the transitions μ_2 do not systematically depend on frame duration.

For Scheme II, the F-ratios given by an ANOVA test show that spatial frequency accounts for 98% of the variance in μ_2 for subject PW (99% for subject JS). Orientation accounts for only 0.4% (PW) and 0.1% (JS).

Comparing Schemes I and II

In the section Motion Computations and their Predictions, we argued that transition invariance was consistent with a single channel motion computation. Werkhoven et al. (1993) have shown that transition invariance is true for a two-dimensional parameter space: spatial frequency and amplitude. Here we have shown that transition invariance approximately (but not entirely) holds for the orientation parameter of the textures involved. Furthermore, transitions are approximately invariant for different frame durations. Therefore, we conclude that the motion computation is dominated by a single motion channel C_0 whose texture grabber (nonlinear preprocessing transformation) is not tuned for texture orientation and largely independent of frame duration. To illustrate this point, we have projected the transition amplitudes which can be represented in a four-dimensional graph—$\mu_i(\omega_v, \alpha_v, \tau)$—on

the μ_i, ω_v-plane. Thus, transition amplitudes μ_i taken from different conditions (different orientations and different frame durations are shown as a function of spatial frequency ω_v). Transitions for Scheme I ($i = 1$) are indicated in Fig. 8 by open circles, those for Scheme II ($i = 2$) by solid circles.

The variance in the data for a particular frequency indicates the range of effects of the two other parameters (orientation and frame duration), which is quite small for the higher frequencies.

The small but consistent violations of transition invariance reported here do, however, indicate that the motion computation is more complex than a single-channel system. When the patches of texture have orthogonal orientations (see right columns of Figs 5, 6 and 7), the transitions are slightly but markedly different for Schemes I and II (summarized in Fig. 8). The fact that transition invariance is not valid for the orientation parameter, suggests that the processing of texture-defined motion receives a modest contribution from additional channels (other than C_0) whose texture grabbers are specific for texture orientation.

To estimate the contribution of multiple channels to motion strength, we will use the following approach. First, we use the observation that the data are well described by a dominant contribution of a single motion channel C_0 with an isotropic texture grabber together with a small contribution of motion channels that have nonisotropic texture grabbers. This contribution of motion channels with nonisotropic texture grabbers is called ϵ. This nonisotropic contribution ϵ will enter the motion computation when the textures in a motion path have orientation differences. When the textures have identical orientation, we find transition invariance and ϵ is negligibly different from zero (Werkhoven et al., 1993).

Second, because the data show that orientation differences systematically weaken the strength of the heterogeneous path, we say that the contribution ϵ is positive

TABLE 3. Relative contribution ϵ/T_s^2 (in %) of motion channels with nonisotropic texture grabbers for observer PW

	$\omega_v = 1.2$ c/deg		$\omega_v = 4.9$ c/deg		$\omega_v = 9.9$ c/deg	
	$\alpha_v = 90$ deg	$\alpha_v = 0$ deg	$\alpha_v = 90$ deg	$\alpha_v = 0$ deg	$\alpha_v = 90$ deg	$\alpha_v = 0$ deg
$\tau = 67$ msec	8	26	0	11	5	8
$\tau = 133$ msec	13	33	2	12	5	6
$\tau = 267$ msec	13	28	0	8	1	5

in the direction of the homogeneous motion path. Since, the anisotropic contribution ϵ will depend mainly on the orientation differences of the textures in the heterogeneous path, it will be similar in strength for competition Schemes I and II, given two textures s and v.

The net motion strength D_1 for Scheme I is determined by standard motion analysis applied to the dominant activity transformation T and the contribution ϵ of other channels that adds positively to the homogeneous path:

$$D_1 = T_s(T_v - T_s) - \epsilon. \tag{12}$$

Similarly, we find the net motion strength D_2 for Scheme II:

$$D_2 = T_v(T_s - T_v) - \epsilon. \tag{13}$$

To proceed, we propose that the activity of a texture grabber is linear in the amplitude of a texture, and that the dependency of T on amplitude and spatial features such as frequency ω and orientation α are separable:

$$T_i(m_i, \omega_i, \alpha_i, \tau) = m_i f(\omega_i, \alpha_i). \tag{14}$$

Separability of spatial frequency and amplitude is supported by the results presented in Werkhoven et al. (1993).

We will focus on the transitions to derive a useful expression for the contribution ϵ. At transition amplitude $m_v = \mu_1$ the motion paths are balanced and $D_1 = 0$. After substitution of equation (14) into equation (12), we find the following expression at transition amplitude for Scheme I:

$$\mu_1 f(\omega_v, \alpha_v) = m_s f(\omega_s, \alpha_s) + \epsilon[m_s f(\omega_s, \alpha_s)]^{-1}. \tag{15}$$

At transition amplitude $m_v = \mu_2$ with Scheme II, we find:

$$\mu_2 f(\omega_v, \alpha_v) = m_s f(\omega_s, \alpha_s)$$
$$[1 + \sqrt{1 - 4\epsilon/(m_s f(\omega_s, \alpha_s))^2}]/2 \tag{16}$$

which reduces to a similar expression as equation (15) using the approximation that ϵ is much smaller than T_s^2:

$$\mu_2 f(\omega_v, \alpha_v) \approx m_s f(\omega_s, \alpha_s) - \epsilon[m_s f(\omega_s, \alpha_s)]^{-1}. \tag{17}$$

The expressions equations (15) and (17) show that the

transition amplitudes for Schemes I and II shift apart symmetrically in case of a small similarity contribution ϵ. This approximation allows an easy derivation of a quantitative measure of a similarity effect.

Equations (15) and (17) can be combined to find an expression for the contribution ϵ. However, because the experiments give us the *relative* strength of motion paths, ϵ can only be measured relative to some reference. We will use the strength of the homogeneous reference path for Scheme I (containing textures s only) as the standard to compare to ϵ. The strength of this standard path is T_s^2. Simple combination of equations (15) and (17) yields:

$$\epsilon/T_s^2 = (\mu_1 - \mu_2)/(\mu_1 + \mu_2). \tag{18}$$

The approximation $\epsilon \ll T_s^2$ provides a useful rule of thumb in the form of equation (18). To quantify the similarity effect in our experiments, however, we combine expressions (15) and (16) to find the general expression for ϵ as a function of μ_1 and μ_2 (allowing for larger ϵ). Furthermore, we say that the activity of textures s equals the average activity of texture v at the transition amplitudes in Schemes I and II: $T_s = f(\omega_v, \alpha_v)(\mu_1 + \mu_2/2$ (recall the symmetry in the shift of the transitions). The relative similarity contribution ϵ with respect to the average motion energy T_s^2 becomes:

$$\epsilon/T_s^2 = 2\mu_2(\mu_1 - 3\mu_2 + \sqrt{\mu_1^2 - 2\mu_1\mu_2 + 5\mu_2^2})/(\mu_1 + \mu_2)^2. \tag{19}$$

In Tables 3 and 4, we presented the relative contribution ϵ as expressed in equation (19) in percentage of the standard strength for all conditions examined here. The contributions of multiple channels are smaller than 14% for all conditions and subjects except for the $\omega_v = 1.2$ c/deg conditions. When the spatial frequency of texture v is 2 octaves lower than that of s ($\omega_v = 1.2$ deg/sec), and the orientation of s and v are orthogonal, the contribution becomes markedly larger. For this "orthogonal" condition, nonisotropic motion channels contribute an average net motion strength that is maximally 33% of the standard strength vs 13% for the "parallel" condition (for subject PW). For subject JS

TABLE 4. Relative contribution ϵ/T_s^2 (in %) of motion channels with nonisotropic texture grabbers for observer JS

	$\omega_v = 1.2$ c/deg		$\omega_v = 4.9$ c/deg		$\omega_v = 9.9$ c/deg	
	$\alpha_v = 90$ deg	$\alpha_v = 0$ deg	$\alpha_v = 90$ deg	$\alpha_v = 0$ deg	$\alpha_v = 90$ deg	$\alpha_v = 0$ deg
$\tau = 67$ msec	21	12	0	7	4	11
$\tau = 133$ msec	22	26	11	9	3	14
$\tau = 267$ msec	32	30	1	12	0	10

we see marked contributions (although smaller than 32%) of a similarity channel for both the "parallel" and "orthogonal" condition when $\omega_v = 1.2$ c/deg. This finding suggests that the possible nonisotropic channels involved in motion-from-texture perception are tuned to low spatial frequencies.

GENERAL DISCUSSION

We have used a special ambiguous motion competition scheme and the notion of transition invariance to reveal the dimensionality and metric of the motion-from-texture extraction process under study. This powerful technique was previous study (Werkhoven et al., 1993) to demonstrate that the motion-from-texture computation is one-dimensional with respect to a two-parameter family of textures (spatial frequency and amplitude). Consequently, the motion computation consists of standard motion analysis applied to a *single* nonlinear transformation of the stimulus input: the *activity transformation*. The *similarity* in spatial structure between elements in a motion path does not determine the motion strength.

Perhaps surprisingly, the study presented here shows that the motion-from-texture computation is also quite insensitive with respect to orientation differences between gratings in a motion path. This means that the activity transformation that precedes standard motion analysis is roughly *isotropic*. Furthermore, the performance of the motion-from-texture mechanism seems invariant for the temporal frequency of the stimulus (between 0.94 and 3.75 Hz). Thus, the assumption of a single channel is *almost* sufficient to predict the strength of second-order motion for a larger family of textures (*orientation*, in addition to spatial frequency and amplitude). To explain the small violations of transition invariance, however, we assume a modest contribution of a secondary channel characterized by non-isotropic preprocessing. The contribution of a channel with non-isotropic preprocessing is 32% of the strength of the dominant isotropic channel.

Texture orientation

We find that a motion path containing differently oriented textures ($\alpha_s = 90$ deg, $\alpha_v = 0$ deg) yields a slightly weaker motion strength than a path containing identically oriented textures ($\alpha_s = \alpha_v = 90$ deg). Recall that a patch of orthogonally oriented texture ($\alpha = 90$ deg) has an orientation of 90 deg relative to the tangent to the annulus of the center but an orientation of 78.75 deg relative to its assigned motion direction. Similarly, a texture oriented parallel to its tangent ($\alpha = 0$ deg) has an orientation of 11.25 deg relative to its assigned motion direction. Due to the circular display geometry, the actual orientation differences between successive textures in a *homogeneous* motion path were 22.5 deg whereas successive patches in the *heterogeneous* paths differed 67.5 deg in orientation.

The weaker motion when $\alpha_v = 0$ deg may be due to (a) the orientation difference $\alpha_v - \alpha_s$ between texture v

and texture s (which we call a correspondence measure) or simply (b) α_v itself: the difference between the orientation of texture v and the motion direction. To distinguish between (a) and (b), one needs to vary α_v and $\alpha_v - \alpha_s$ independently for a given α_s, which is not possible with the present configuration. Since we cannot rule out explanation (b), the small deviations of transition invariance are not necessarily due to correspondence measures. Werkhoven et al. (1990a) reported on a method to solve this question for oriented elements in the first-order motion domain and showed that the relative orientation of elements in a motion path are irrelevant to motion strength. Instead, the element orientation relative to the *motion direction* rules first-order motion strength. We have not yet applied this method to the second-order motion domain.

Stimulus display rate

A second important finding is that the parameters extracted by motion computation proposed here are invariant for frame durations between 67 and 267 msec (corresponding to stimulus temporal frequencies between 3.75 and 0.94 Hz). As argued in the Introduction, one might expect an increasing contribution of more complex (and perhaps more time consuming) correspondence measurements as the presentation time for each increases. Temporal frequencies of 3.75 Hz (together with the small displacement of 37 min arc) would be classified under low-level or short-range motion processing, as opposed to *long-range* motion processing (Anstis, 1970). However, even when the presentation time of each patch of texture in a motion path is 267 msec (0.94 Hz), similarities between textures do not seem to play a role in motion processing. For shorter frame durations, 67 msec, we find transitions equal to those for 267 msec. Thus, we find no shift towards different motion computations at longer frame durations.

Orientation and frequency selectivity in first- and second-order motion perception

In this paper we have argued that motion-from-texture processing can be understood by assuming a dominant channel consisting of isotropic low-pass preprocessing followed by motion energy analysis. To explain the small similarity effects found for texture orientation at low spatial frequencies we have assumed a second weak nonisotropic channel tuned to low spatial frequencies. Thus, motion-from-texture analysis seems to be an extremely low-dimensional process.

Interestingly, first-order motion analysis is believed to be composed of multiple channels that differ in their spatial and temporal frequency tuning and their orientation selectivity. Anderson and Burr (1985), for example, employed a masking technique (measuring the sensitivity to a sinewave grating with and without a superimposed high contrast mask grating) to reveal the involvement of multiple channels tuned to different spatial frequencies and orientations (see also Anderson, Burr & Morrone, 1991). Studies on motion aftereffects (e.g. Cameron, Baker & Boulton, 1992) provide further

support for a multiple spatial frequency channel model of first-order motion processing.

Before discussing any possible discrepancies between spatial frequency and orientation selectivity in first- and second-order motion processing, it is helpful to realize the following. Generally, the selectivity for spatial frequency and orientation in *first*-order motion processes is attributed to the spatial properties of the subunits of bilocal motion detectors. These subunits filter the input in both space and time and are thought to be followed by a nonlinear operation to compute, e.g. motion energy. The receptive fields of these subunits have a fine structure that determines the selectivity of a bilocal motion detector to the orientation and spatial frequency of a moving grating (see e.g. Anderson *et al.*, 1991).

The selectivity to orientation and spatial frequency discussed in this paper, however, should be attributed to the properties of the preprocessing stage that transforms the moving texture-modulation into a moving activity-modulation. This moving activity-modulation can be sensed by a bilocal motion detector. Obviously, the properties of the subunits of the motion detector can account for the effects of the orientation and frequency of the *activity-modulation*. Conversely, the properties of the preprocessing stage account for the effects of the orientation and frequency of the *textures*.

Our stimuli are basically one-dimensional moving modulations of texture type and we have examined second-order motion strength as a function of textural properties. We have not yet studied the dependence of second-order motion strength on the properties of the modulation function itself. Such a study would be necessary to compare orientation and frequency selectivity in first- and second-order processing.

CONCLUSION

We have shown that the strength of texture defined motion is dominated by a single nonlinear transformation followed by standard motion analysis for a family of textures that includes orientation, spatial frequency and amplitude. Given the relative insensitivity of motion-from-texture processing to correspondences between these simple features of textures in a motion path, it is unlikely that more detailed aspects of spatial fine

structure will trigger other more elaborate motion-from-texture mechanisms.

Our findings decisively refute the historically influential notion that shape and contrast correspondence underlie the motion-from-texture computation.

REFERENCES

Anderson, S. J. & Burr, D. C. (1985). Spatial and temporal selectivity of the human motion detection system. *Vision Research, 25,* 1147–1154.

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

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research, 10,* 1411–1430.

Balliet, R. & Nakayama, K. (1978). Training of voluntary torsion. *Investigative Ophthalmology & Visual Sciences, 17,* 303–314.

Cameron, E. L., Baker, C. L. Jr & Boulton, J. C. (1992). Spatial frequency selective mechanisms underlying the motion aftereffect. *Vision Research, 32,* 561–568.

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. (1991). Texture quilts: Basic tools for studying motion from texture. *Journal of Mathematical Psychology, 35,* 411–442.

Farrell, J. E., Pavel, M. & Sperling, G. (1990). The visible persistence of stimuli in stroboscopic motion. *Vision Research, 29,* 1511–1523.

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

Pantle, A. & Picciano, L. (1976). A multistable movement display: Evidence for two separate motion systems in human vision. *Science, 193,* 500–502.

van Santeen, J. P. H. & Sperling, G. (1984) Temporal covariance model of human motion perception. *Journal of the Optical Society of America A, 1,* 451–473.

Watson, A. B. (1986). Apparent motion occurs only between similar spatial frequencies. *Vision Research, 26,* 1727–1730.

Werkhoven, P., Snippe, H. P. & Koenderink, J. J. (1990a). Effects of element orientation on apparent motion perception. *Perception & Psychophysics, 47,* 509–525.

Werkhoven, P., Snippe, H. P. & Koenderink, J. J. (1990b). Metrics for the strength of low level motion perception. *Journal of Visual Communication and Image Representation, 1,* 176–188.

Werkhoven, P., Sperling, G. & Chubb, C. (1993). The dimensionality of texture defined motion: A single channel theory. *Vision Research, 33,* 463–485.