1992

Visual Neuroscience (1992), **9**, 105–123. Printed in the USA.

Evoked potential and psychophysical analysis of Fourier and non-Fourier motion mechanisms

JONATHAN D. VICTOR[1] AND MARY M. CONTE[2]

[1]Department of Neurology and Neuroscience, Cornell University Medical College, New York, New York
[2]Laboratory of Biophysics, The Rockefeller University, New York, New York

(RECEIVED September 11, 1991; ACCEPTED December 2, 1991)

Abstract

Some visual stimuli produce a strong percept of motion, even though they fail to excite motion detectors based on Fourier energy or cross correlation. Models which suffice to explain the motion percept in these non-Fourier motion (NFM) stimuli include linear spatiotemporal filtering, followed by rectification, followed by standard motion analysis (Chubb & Sperling 1988). We used the human "motion-onset" evoked potential, which has been assigned to area 17 on the basis of work in the macaque (van Dijk et al., 1986; van Dijk & Spekreijse, 1989), to investigate the neural substrate of the processing stages postulated in the above models. Motion-onset VEPs elicited by FM and NFM matched for spatial and temporal characteristics were indistinguishable in temporal characteristics and scalp topography at a transverse chain of electrodes. Addition of textural cues (granularity and higher-order form) did not influence the response dynamics or scalp topography of NFM responses. However, comparison of responses to NFM stimuli and related stimuli without coherent motion but similar spatial and temporal properties showed that the motion-onset responses were distinct from responses to the onset of fixed flicker-defined contours not undergoing coherent motion. We discuss the implications of these results for computational models of motion analysis.

Keywords: Human, Modeling, Motion mechanisms, Nonlinear interactions, Rectification, Visual-evoked potentials

Introduction

The retinal image of real moving objects can be mimicked by images of sources in which no movement is present (such as a movie or a CRT display). That is, the percept of motion is not an intrinsic aspect of the spatiotemporal pattern of light which falls on the retina; rather, motion requires a computation for its extraction. A variety of proposals for how this computation might take place have been advanced (Nakayama, 1985), including methods based on cross correlation (Reichardt, 1961), spatiotemporal motion energy (Adelson & Bergen, 1985; Heeger, 1987), and detection of spatiotemporal gradients. These computational schemes are closely related (van Santen & Sperling, 1985; Nakayama, 1985; Simoncelli & Adelson, 1991). We will call motion signals extracted by these computations Fourier motion (FM), since the motion signals are manifest in the spatiotemporal Fourier components of the stimulus. Equivalently, stimuli which elicit FM in the direction $\mathbf{v}$ contain bands of correlation along lines $\mathbf{x} = \mathbf{v}t$ in space-time.

Although FM models suffice to account for the perception of motion as elicited by many natural and artificial scenes, they do not account for the perception of motion in all cases, such as the drift-balanced stimuli introduced by Chubb and Sperling (1988). Motion percepts which cannot be accounted for by the above standard models will be called non-Fourier motion (NFM). A simple example of an NFM stimulus is a flickering bar moving across a static random background of equal mean luminance.

One way of accounting for NFM is to postulate an initial stage of rectification (perhaps along with local spatiotemporal filtering) prior to motion detection *via* one of the computational schemes for FM mentioned above. Alternatively, the percept of motion in NFM textures might be driven by an entirely separate kind of computation, whose relationship to standard motion detection might be analogous to the relationship between short-range and long-range motion (Braddick, 1974, 1980).

One approach to separate these possibilities is suggested by the work of van Dijk et al. (van Dijk et al., 1986; van Dijk & Spekreijse, 1989), who found that onset of FM stimuli generates a robust visual-evoked potential (VEP) in man and macaque, and that (in the macaque) this response is generated by area 17. We therefore asked whether the onset of NFM would elicit a similar VEP, and whether this VEP was analogous to the VEP elicited by FM in its dependence on contrast and velocity. As shown below, onset of NFM indeed generates a response which is very similar to the onset of FM. Since NFM stimuli allow independent manipulation of motion and contrast, we then used NFM stimuli to determine to what extent this motion-onset

Reprint requests to: Jonathan D. Victor, Department of Neurology and Neuroscience, Cornell University Medical College, 1300 York Avenue, New York City, NY 10021, USA.

VEP was associated with motion as such, rather than spatiotemporal luminance changes not associated with the percept of smooth motion. Finally, the close correspondence of VEP responses elicited by FM and NFM stimuli places strong constraints on computational models.

Methods

Visual stimuli

The stimuli in these experiments are variations on a common design. We set out the common design and the basic stimulus in detail, and indicate how variations on this design were made. Each frame of the stimulus consisted of 16 vertical strips 1.1 deg across, as diagrammed in Fig. 1A. Each vertical strip was 256 checks high and 16 checks across. The checks subtended 4.125 min at the viewing distance of 28.5 cm. Each check consisted of one pixel of the raster display (see below). The coloring schemes for these strips consisted of a variety of rules, to allow independent manipulation of cues to motion (Fourier and non-Fourier) and cues to strip boundaries (flicker, spatial frequency, and local form elements). The positions of the strips on the display changed from frame to frame.

The details of how the strips are colored for the basic FM stimulus are shown in Fig. 1B, which consists of an enlargement of a fixed region of the stimulus on each of five successive frames. In the first frame, check luminance values are independently assigned with probability 0.5 to dark or light. In successive frames, a single vertical column of checks (set off by heavy lines for illustrative purposes in Fig. 1B) is translated by a single column to the left. This occludes its neighbor to the left, and reveals a fresh column of random checks in its former position. As this column undergoes successive translations, it continues to occlude successive columns of the original background, and the checks that it reveals are always that of the original background. Thus, the percept elicited by a region of this stimulus is that of a thin vertical column of random checks moving in front of a background of static checks. The full stimulus consists of 16 such independently colored columns, spaced by 1.1 deg.

This stimulus will be denoted *R/R*-drift, to indicate that all strips are randomly colored and that their border drifts smoothly. Since the checks of the moving strip constitute a band of correlations in space-time, they are an example of FM.

To convert this stimulus into a NFM stimulus, the checks in the moving vertical strip are recolored randomly each time the strip is translated laterally (Fig. 1C). The percept is that of a randomly flickering vertical column of random checks moving in front of a background of static checks. Recoloring these checks with each lateral translation removes the band of correlations in space-time otherwise associated with the strip in the FM stimulus described above, and results in a NFM stimulus. We denote this stimulus *R/R*-drift*, to indicate that the stimulus is identical to *R/R*-drift except that the border strip is recolored with each lateral translation.

For both *R/R*-drift and *R/R*-drift*, the "standard" drift velocity consisted of translation of the vertical strip by one check (4.125 min) every eight refreshes (29.6 ms), which provided a drift rate of $v_0 = 2.32$ deg/s. More rapid velocities were realized by translations of 2, 4, or 8 checks, providing velocities from $2v_0 = 4.64$ to $8v_0 = 18.6$ deg/s. This ensured that the

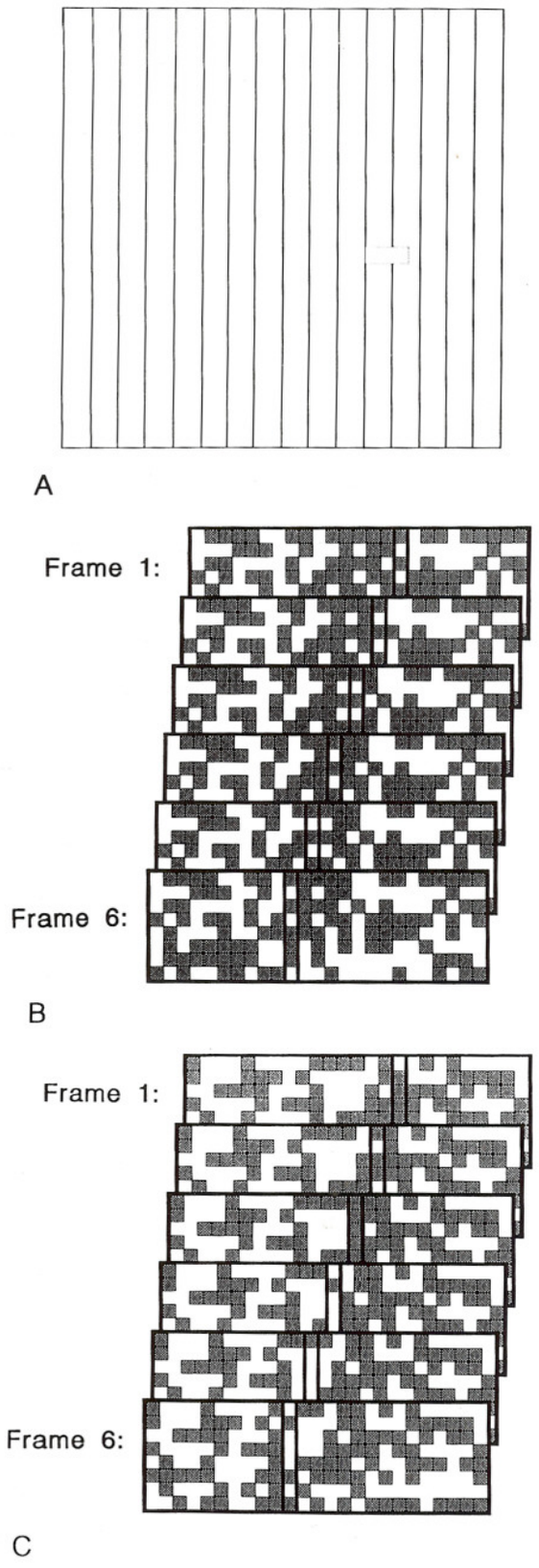

Fig. 1. A: Overall scheme of motion stimuli. B: Detail of a small region of the *R/R*-drift stimulus. This is a FM stimulus. C: Detail of a small region of the *R/R*-drift* stimulus. This is an NFM stimulus.

temporal characteristics of the stimulus did not change as velocity increased. Note that at the fastest velocity used, the direction of motion is ambiguous: each new bar position is halfway in between two previous positions. Slower velocities ($\frac{1}{2}v_0 = 1.61$ deg/s and $\frac{1}{4}v_0 = 0.58$ deg/s) were realized by translation by one check every 16 or 32 refreshes.

Use as evoked-potential stimuli

For evoked-potential studies, we alternated epochs of motion (L or R) with epochs in which the vertical strips were immobile (I). Each such epoch lasted $T_d = 473$ ms (128 refreshes). Epochs with leftward and rightward motion were balanced, either according to a simple periodic scheme (LIRI) or longer schemes (LILIRILIRILIRIRI and LILILIRILIRIRIRI).

All stimulus cycles thus consisted of epochs of length T_d of motion, followed by epochs of length T_d of no motion. In initial experiments, we averaged responses with respect to the period of the complete sequence ($4T_d$ for LIRI; $16T_d$ for the two longer schemes). We found that averaged responses to the L, R, and I epochs did not depend on which sequence was used. Therefore, we subsequently averaged VEPs over a period of length $2T_d$ for all schemes. This average contained responses to leftward and rightward drift superimposed. To separate responses to leftward and rightward drift, we retained the average of responses to the scheme LIRI with respect to a period of length $4T_d$. This averaged response was then reprocessed two ways: (1) the first and second halves of the averaged response were in turn averaged, and (2) half of the difference of the first and second halves of the averaged response (first half minus second half) was computed. The first maneuver merely recovers the average with respect to the period of length $2T_d$, in which responses to leftward and rightward motion are superimposed. The second maneuver isolates the difference in the responses to leftward and rightward drift.

Stimulus variations

We also used two sets of variations on these stimuli. In the first set of variations, we colored the vertical strips (and not just their boundaries) according to one of several rules to examine the relevance of spatial-frequency content and pattern. Three rules were used: *r*andom coloring (as in the basic stimulus), which we denote by *R*; coloring with *c*oarse texture elements (random 1 × 4 vertical rectangles), which we will denote by *C*, and coloring with the *e*ven texture of Julesz et al. (1978), which we will denote by *E*. All of these colorings assign half of the checks to each of two brightnesses. The *C* coloring and the *R* coloring differ in spatial-frequency content. The *E* coloring and the *R* coloring are preattentively discriminable but do not differ in spatial-frequency content. For example, in the "even-random" (*E*/*R*-band) stimulus, alternate strips were colored according to the rule for the even texture or at random. We use the designation band to indicate that individual frames of these stimuli were segregated into bands by their colorings (the above rules), and not just by their drifting boundaries.

In all cases, motion at the standard velocity $v_0 = 2.32$ deg/s was introduced by shifting the position of the borders between vertical strips by a single check every 29.6 ms frame (eight refreshes). The checks assigned to the boundaries between vertical strips were recolored to match the new rule to which they are assigned. This is illustrated for the *E*/*R*-band stimulus in Fig. 2A. Additionally, we also used stimuli in which checks on the interior of the vertical strips (and not just those on the boundary) were recolored every frame. For example, in the E^*/R-band stimulus (Fig. 2B), every even strip was recolored on each frame, but checks in the random strips were recolored only as the boundaries moved. Similarly, the stimulus E/R^*-band denotes a stimulus in which all checks within the random strips are recolored on every frame, but only the changing boundaries of the even strips are recolored. The stimulus E^*/R^*-band denotes one in which all checks are recolored on each frame. Analogous stimuli were constructed based on the coloring rule *C* (Fig. 2C shows C/R^*-band) and with all bands assigned the coloring rule *R* (Fig. 2D shows *R*/*R*-band).

The percepts elicited by these stimuli are that of moving partitions between textures. All of these stimuli are examples of NFM, since the new coloring on each frame is uncorrelated (at second order) with current or previous colorings of its neighbors. This is a key difference between the stimulus *R*/*R*-band and the stimulus *R*/*R*-drift, which is an example of FM.

Some of these stimuli are identical to the non-Fourier motion stimuli we have used previously in psychophysical studies (Victor & Conte, 1990): *E*/*R*-band = *FH*, E^*/R^*-band = *H*, *C*/*R*-band = *FG*, C^*/R^*-band = *G*, and *R*/*R*-band = *F*. We use the new notation here to facilitate the description of additional variations on these stimuli. For this set of stimulus variations, the layout of 16 vertical strips each 16 checks wide was modified to that of eight vertical strips each 32 checks wide, to match the parameters of the previous study.

The second set of variations applies only to the NFM stimulus *R*/*R*-drift*, and was used to compare responses to drift with responses to stimuli with similar spatiotemporal characteristics but no coherent motion. In the stimulus denoted *R*/*R*-fixed* (Fig. 3A), the column of checks that made up the strip boundaries were recolored with each frame, but their position was not shifted. (The analogous variation on the FM stimulus *R*/*R*-drift of Fig. 1B would simply be a static display of randomly colored checks, since neither the position nor the coloring of the vertical bar would change. We did not use this stimulus.) Additionally, we allowed the position of the column of checks that formed the strip boundaries to be randomly repositioned and recolored each frame, providing a stimulus denoted *R*/*R*-jitter* (Fig. 3B).

General experimental procedures

The stimuli described above were produced on a Tektronix 608 monitor. Control signals for the 270.3 Hz raster display [horizontal (X), vertical (Y), and intensity (Z)] were generated by specialized electronics (Milkman et al., 1980) interfaced to a DEC 11/73 computer. The display subtended 17.6 deg and had a luminance of 150 cd/m^2. Stimuli had a contrast $[(I_{max} - I_{min})/(I_{max} + I_{min})]$ of 0.4, and were viewed binocularly at a distance of 28.5 cm. Each pixel of the 256 × 256-pixel display subtended 4.125 min. A 4-min dark fixation point was placed in the center of the display.

The subject pool for VEP and psychophysical studies consisted of three normal volunteers, one male and two female; one of these subjects was naive to the purpose of the experiments. Subjects ranged in age from 30 to 41 years, and were corrected to normal visual acuity if necessary.

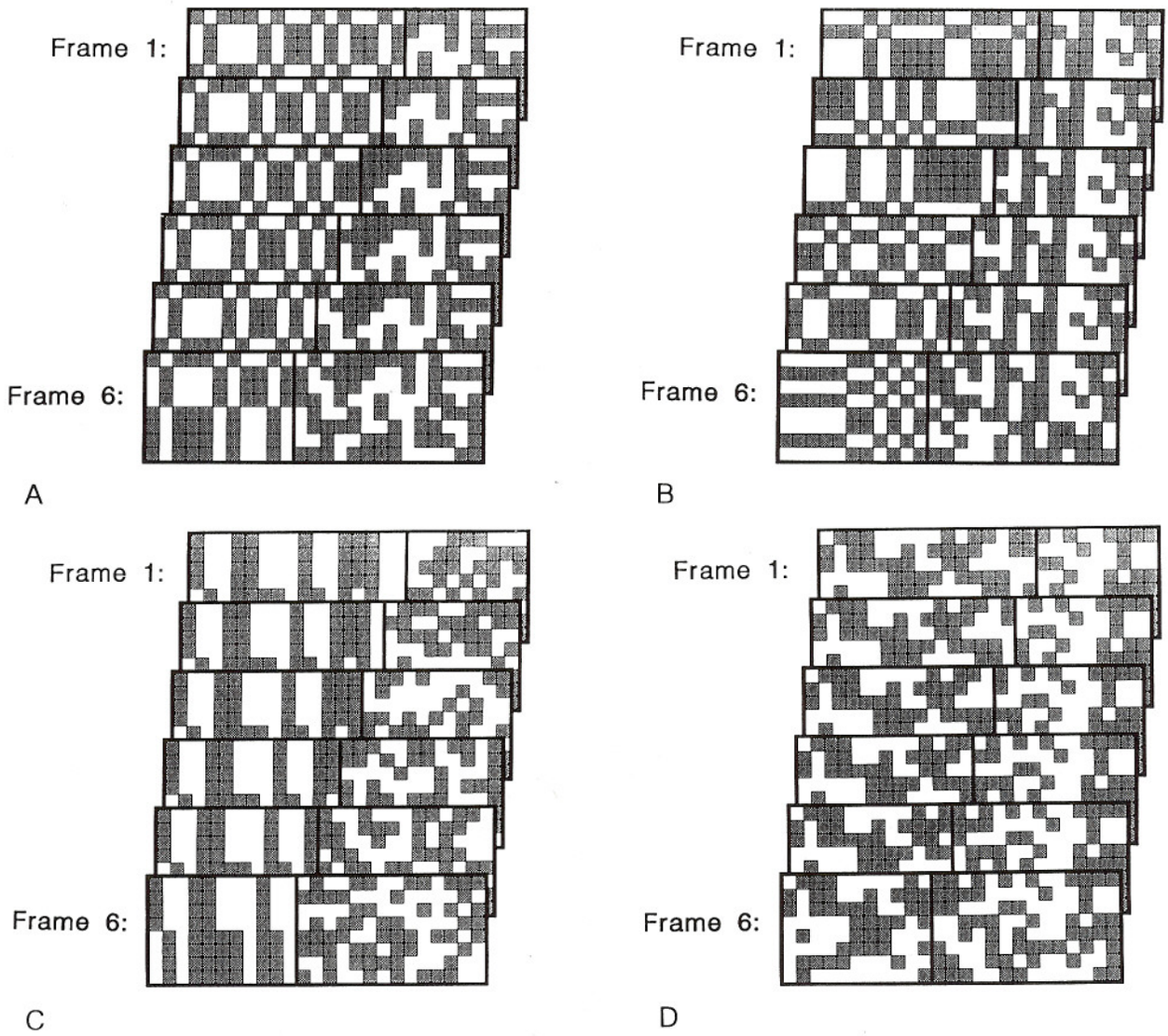

Fig. 2. A: Detail of a small region of the *E/R*-band stimulus. B: Detail of a small region of the *E*/R*-band stimulus. C: Detail of a small region of the *C/R**-band stimulus. D: Detail of a small region of the *R/R*-band stimulus. These are all NFM stimuli. See Fig. 1A for overall stimulus layout.

Measurement of visual-evoked potentials

Brain electrical signals were recorded with standard gold cup electrodes applied to the scalp with electrolyte paste. Input impedances were typically less than 5000 Ω.

In studies in which scalp topography was of interest, recordings were made at a chain of seven electrodes (L_{75}, L_{50}, L_{25}, O_z, R_{25}, R_{50}, and R_{75}) referenced to C_z. Here, L_{75} denotes a position 7.5 cm lateral to O_z on the left, etc. In the remaining studies, recordings were made at C_z-O_z alone. Signals were amplified 10,000-fold and bandpass filtered (0.03–100 Hz) prior to digitization at the raster refresh rate of 270.3 Hz. Artifact rejection was performed by the computer prior to averaging.

In one set of control experiments, electrooculographic (EOG) records were made with gold cup electrodes applied at the outer canthi. EOG signals were processed in the same way as the EEG signals.

The averaged VEPs were saved by the 11/73 computer for offline analysis, which included Fourier analysis, modeling, and factor analysis. Each raw waveform represents an average of responses recorded in at least three 1-min runs, presented in randomized blocks. Recording sessions were limited to 2 h, with breaks as needed to maintain subject alertness. All responses are displayed with positivity at the occipital electrode plotted downward.

Psychophysical methods

The observer's forced-choice task was to distinguish leftward drift from rightward drift. Trials were presented in blocks of 20. Before data collection, observers were given practice (with feedback) with free-view and timed presentations of examples of the target stimuli. Practice was allowed until performance stabilized; no feedback was given during data collection. A set of two or more blocks were presented in random order once performance had stabilized.

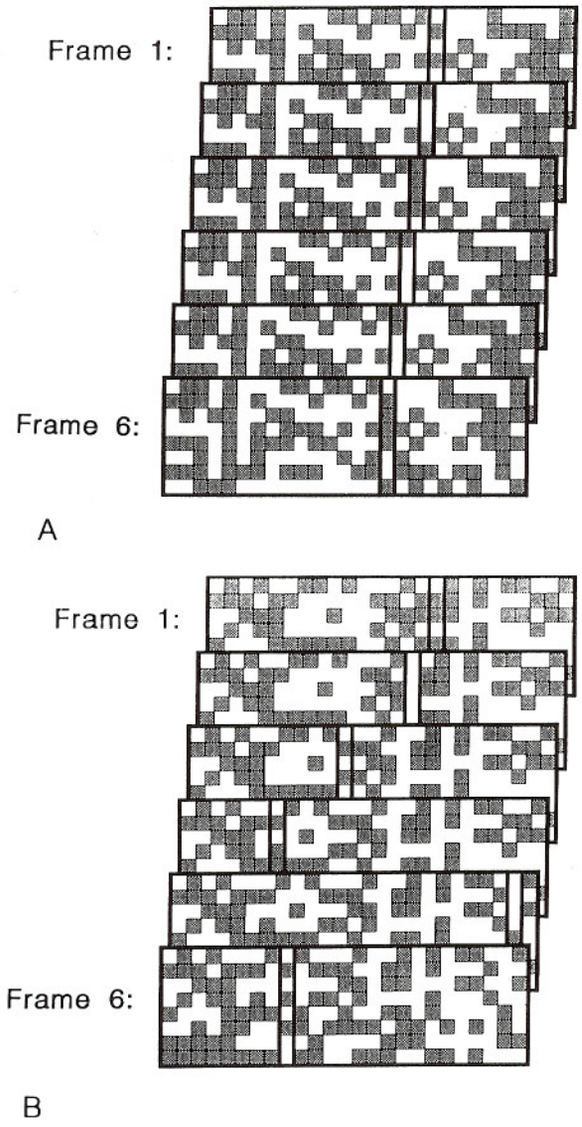

Fig. 3. A: Detail of a small region of the *R/R*-fixed* stimulus. B: Detail of a small region of the *R/R*-jitter* stimulus. See Fig. 1A for overall stimulus layout.

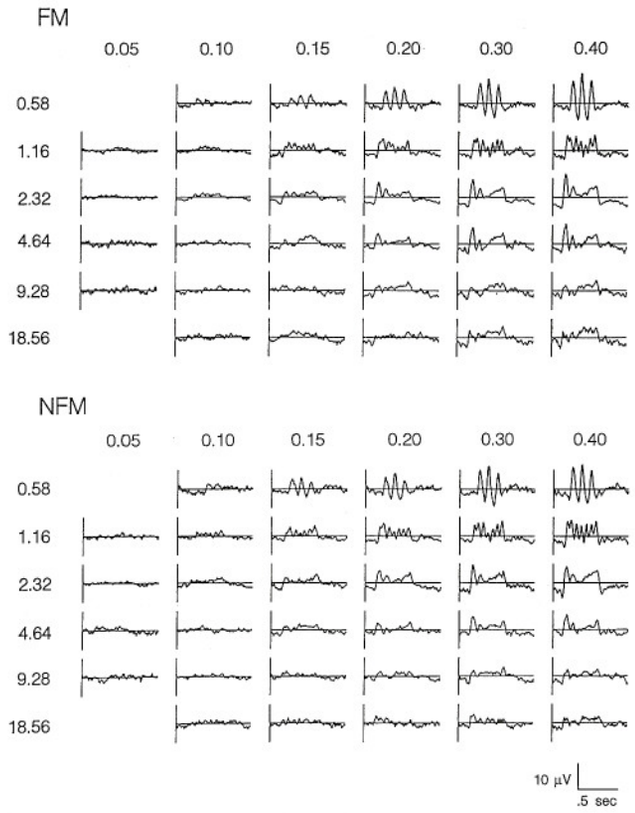

Fig. 4. VEPs elicited by FM (*R/R*-drift) and NFM (*R/R*-drift*) stimuli. Onset of motion occurs at the beginning of each histogram; cessation of motion occurs in the middle of each histogram. Each histogram represents an average of at least 256 epochs of drift, with leftward and rightward drift superimposed. Derivation: C_z-O_z. Occiput positivity is plotted downward. Subject: MC.

Results

VEP responses to Fourier and non-Fourier motion stimuli

Fig. 4 compares responses to onset of FM and NFM, over a range of contrasts (0.05–0.4) and velocities (0.58–18.6 deg/s). The FM stimulus was *R/R*-drift; the NFM stimulus was *R/R*-drift*. Responses generally consisted of an occiput-negative transient followed by a more sustained component. For the two lowest velocities, the sustained component is fragmented into individual responses to each translation of the stimulus (118.4-ms intervals for 0.58 deg/s and 59.2 ms for 1.16 deg/s). At the lowest contrast at which responses are clearly distinguishable from noise (0.1), a velocity tuning is evident, with a maximal response at or near 2.32 deg/s. At higher contrasts, the prominence of the responses to individual shifts at low velocities makes it difficult to gauge velocity tuning from the visual inspection of the waveforms, but a velocity tuning is evident upon Fourier analysis (see below). These features are present for both FM and NFM responses, and the overall similarity of the waveforms across these conditions is striking.

To further analyze the responses and to estimate any differences between responses to FM and NFM conditions, we calculated the Fourier components of the response with respect to the averaging period $2T_d$. While Fourier components up to the eighth harmonic were significantly different from zero [by the T^2_{circ} statistic (Victor & Mast, 1991)], amplitudes above the fourth harmonic were typically less than 10% of the fundamental amplitudes, and were therefore excluded from further analysis. The overall size of each response was then quantified by the sum of the first four Fourier components.

A summary of the dependence of response size on stimulus contrast (for stimulus velocity 2.32 deg/s) is shown in Fig. 5. For all three subjects, the responses to FM and NFM grow approximately linearly with contrast over the range studied. A summary of the dependence of response size on stimulus velocity (for stimulus contrast 0.4) is shown in Fig. 6. Responses are maximal at 2.32 deg/s, and there are no consistent differences between FM and NFM responses. The velocity tuning of the responses provides evidence that we are not simply measuring responses to appearance and disappearance of flicker—the

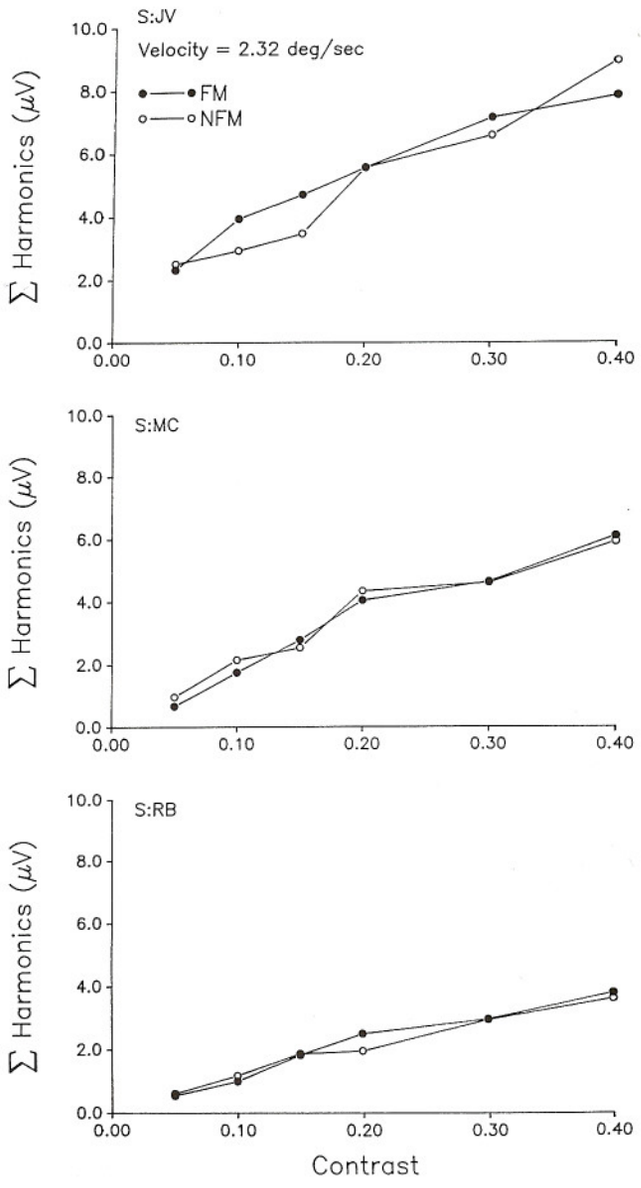

Fig. 5. Dependence of response size (sum of first four Fourier amplitudes) on stimulus contrast, for stimulus velocity 2.32 deg/s.

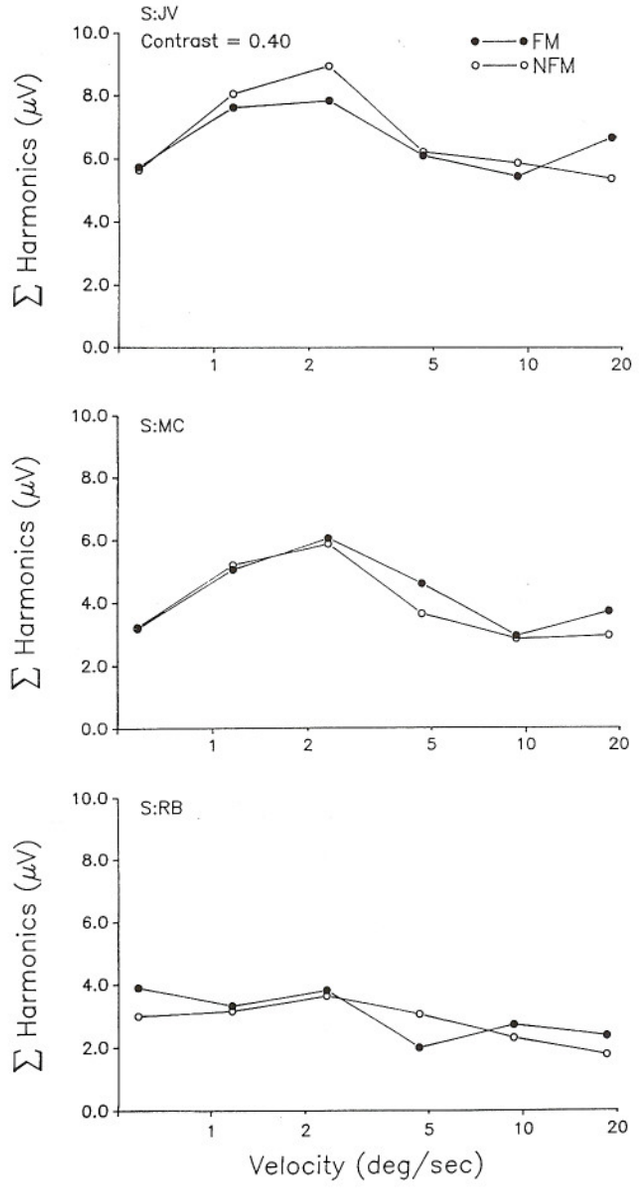

Fig. 6. Dependence of response size (sum of first four Fourier amplitudes) on stimulus velocity, for stimulus contrast 0.4.

temporal characteristics of the stimulus were identical at velocities 2.32 deg/s and above.

We used the T^2_{circ} statistic to assay the statistical significance of the differences between FM and NFM responses. This statistic compares Fourier components of two steady-state responses, in a manner which incorporates both amplitude and phase. Within each subject, there were 34 experimental conditions (six contrasts, six velocities, but the two extreme velocities were not tested at the lowest contrast), and four Fourier components measured for each condition. This yielded a total of $4 \times 34 = 136$ comparisons between FM and NFM responses for each subject. By chance alone, $0.05 \times 136 = 6.8$ comparisons would be expected to be significant at $P = 0.05$. For each of the subjects RB and MC, seven of the comparisons were found to be "significant" at $P = 0.05$. That is, there was no statistical evidence of any difference between the responses to FM and NFM. For subject JV, 20 such comparisons were significant at $P = 0.05$. Seven of these statistically significant differences were for the highest velocity (18.6 deg/s); fundamental responses to NF motion were approximately 30% smaller than responses to FM at contrasts 0.1–0.3. No similar clustering was seen for other harmonics or in the other subjects.

Eye movements

As a control for the possible influence of eye movements, we measured VEPs elicited by FM and NFM stimuli under the usual experimental conditions (subject instructed to fixate a marker in the center of the display) and with the subject instructed to track the motion of the stimulus. An EOG record was used to verify subject compliance (see Methods). The FM stimulus *R/R*-drift and the NFM stimulus *R/R*-drift* were

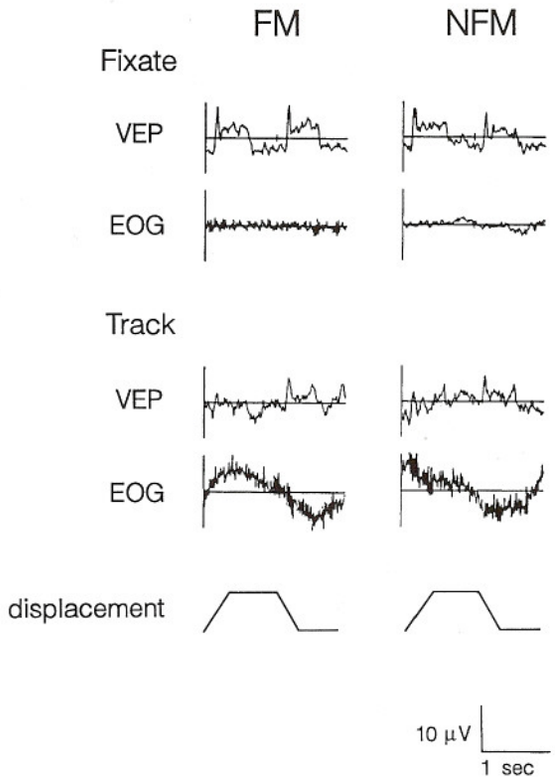

Fig. 7. VEPs elicited by FM (*R/R*-drift) and NFM (*R/R*-drift*) stimuli during conditions of fixation and tracking. Motion trajectory is indicated below the first column. Onset of leftward motion occurs at the beginning of each histogram; onset of rightward motion occurs in the middle of each histogram. Each histogram represents an average of 32 repeats of the stimulus cycle. VEP derivation: C_z-0_z. Occiput positivity is plotted downward. EOG derivation: left eye to right eye; leftward eye deviation is plotted upward. VEP and EOG data have been smoothed by a bell of half-width 5 ms. Subject: RB.

used, at a velocity of 2.32 deg/s and a contrast of 0.4. The stimulus sequence was LIRI. As seen in Fig. 7, when the subject was instructed to track, the eyes moved roughly sinusoidally. The waveforms elicited by either FM or NFM changed dramatically under conditions of tracking. During tracking, the responses contained a transient component at motion onset and at motion cessation, and the sustained component was much less prominent. That is, a transient component was present when there was a change in retinal image motion. During fixation, such changes occur only at stimulus motion onset. But with tracking, such changes also occur when the stimulus ceases motion but the eye is moving at the tracking velocity. This dependence on retinal image motion is evidence that the VEP responses are related to motion, and not just flicker.

Scalp topography of responses

To search for possible differences between VEPs elicited by FM and NFM stimuli, we examined the scalp topography of the waveforms at a transverse chain of derivations from C_z-L_{75} to C_z-R_{75}.

Data from two subjects are presented in Fig. 8; data from the third subject (RB) resembled that of MC. Despite differences between subjects in the shapes of the waveforms, certain features are evident within the dataset from each subject. The striking feature is the detailed correspondence of response waveforms elicited by FM and NFM at all scalp positions, and at all contrast levels.

The responses to motion onset were attenuated at recording sites 5 cm or more away from the midline. However, these responses were not simply a scaled-down version of the responses recorded in the near midline leads. Rather, in these laterally placed electrodes, the size of the later portion of the waveform was attenuated out of proportion to the attenuation of the initial occiput negativity. To analyze this change in response waveform, we applied a principal components analysis in the Fourier domain (Gutowitz et al., 1986). The response $\mathbf{R}(d)$ at derivation d($d = C_z$-$L_{75}, \ldots, C_z$-$O_z, \ldots, C_z$-R_{75}) was considered to be a vector whose elements are its first four Fourier components. That is,

$$\mathbf{R}(d) = (R_1(d), R_2(d), R_3(d), R_4(d)), \tag{1}$$

where $R_n(d)$ is a complex number which represents the amplitude and phase of the nth Fourier component of the waveform as recorded at derivation d. We then represented the responses $\mathbf{R}(d)$ as linear combinations of a small number H of "principal components" $\mathbf{P}_h$, where $\mathbf{P}_h$ is a response vector of the form, eq. (1).

$$\mathbf{R}(d) = \sum_{h=1}^{H} w_h(d)\mathbf{P}_h. \tag{2}$$

That is, the weights $w_h(d)$ indicate the contribution of the principal components $\mathbf{P}_h$ to the response $\mathbf{R}(d)$. The weights w_h are normalized across derivations by

$$\sum_d [w_h(d)]^2 = 1. \tag{3}$$

If all responses were scalar multiples of a common waveform, then the representation eq. (2) would be adequate with $H = 1$. Systematic departures from this simple picture will be manifest by the significance of two (or more) principal components $\mathbf{P}_h$ in the representation, eq. (2).

Part A of Table 1 presents the fraction of the response power accounted for by the representation eq. (2), with 1, 2, and 3 principal components. For both FM and NFM responses, 97–98% of the power is accounted for by the first two principal components. At least 99% of the power is accounted for by the addition of a third principal component. To compare scalp topography of FM and NFM responses, it thus suffices to compare the weights $w_h(d)$ in a representation, eq. (2), consisting of two or three principal components.

Even with two principal components ($H = 2$), the representation by eq. (2) is not unique (Cattel, 1978). The representation determines the H-dimensional subspace which most nearly contains (in a least-squares sense) the responses $\mathbf{R}(d)$. Though this best-fitting subspace is unique, the choice of basis vectors $w_h(d)$ within that subspace is not uniquely determined. A change of basis vectors $w_h(d)$ induces a linear transformation among the principal components $\mathbf{P}_h$; the transformed representation provides an equally adequate fit. Thus, the particular values $w_h(d)$ and the particular waveforms $\mathbf{P}_h$ cannot be regarded as having a direct interpretation. For this reason, we cannot

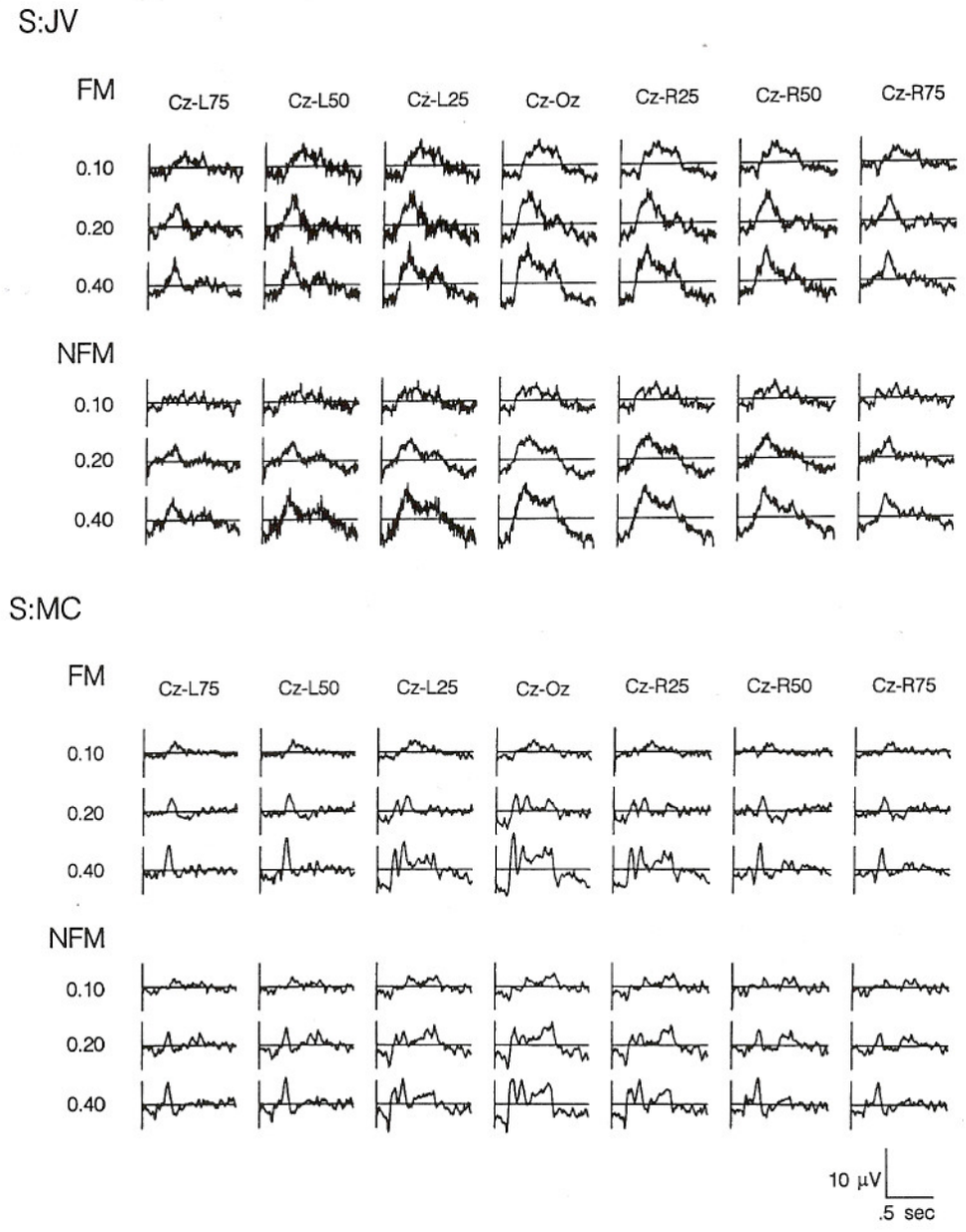

Fig. 8. Scalp topography of VEPs elicited by FM (*R/R*-drift) and NFM (*R/R*-drift*) stimuli. Onset of motion occurs at the beginning of each histogram; cessation of motion occurs in the middle of each histogram. Each histogram represents an average of at least 256 epochs of drift, with leftward and rightward drift superimposed. Subjects: JV and MC.

compare the individual values $w_h(d)$. Rather, we compare algebraic "invariants" constructed from them (Gutowitz et al., 1986). For a representation consisting of two principal components [$H = 2$ in eq. (2)] the invariant $I^{[2]}$ is a function of two derivations (d and d'). $I^{[2]}(d,d')$ is defined by

$$I^{[2]}(d,d') = \det \begin{vmatrix} w_1(d) & w_1(d') \\ w_2(d) & w_2(d') \end{vmatrix}. \tag{4}$$

Analogously, an invariant $I^{[3]}(d,d',d'')$ may be constructed for a representation in terms of three principal components:

$$I^{[3]}(d,d',d'') = \det \begin{vmatrix} w_1(d) & w_1(d') & w_1(d'') \\ w_2(d) & w_2(d') & w_2(d'') \\ w_3(d) & w_3(d') & w_3(d'') \end{vmatrix}. \tag{5}$$

Because the invariants are determinants, their values are independent of a change in basis vectors, up to a common scale factor (Gutowitz et al., 1986). The magnitude of the invariant $I^{[2]}(d,d')$ indicates the extent to which the responses measured at derivations d and d' fail to be scalar multiples of each other. Analogously, the magnitude of the invariant $I^{[3]}(d,d',d'')$ indicates the extent to which the responses measured at derivations d, d', and d'' fail to lie in a common plane. The invariant $I^{[1]}(d)$ is merely the weight $w_1(d)$, since there is no ambiguity in choice of basis vectors for a representation in terms of a single principal component. The invariants $I^{[1]}$, $I^{[2]}$, $I^{[3]}$, ... describe the spatial weights of the principal components expansion, eq. (2), up to the intrinsic ambiguity related to arbitrary linear combinations among the weights.

We recognize that the identification and localization of intracranial sources from scalp potentials is an ill-posed problem. The invariant approach pursued here allows scalp topographies

to be compared across subjects and conditions, without the need to associate principal components with generators, or to solve the source problem. For a fuller discussion of this issue, see Gutowitz et al. (1986).

Fig. 9 shows the first two weight vectors and the invariants $I^{[2]}(d,d')$ derived from the data of Fig. 8 at a contrast of 0.4. Qualitatively, the first principal component corresponds to the overall decrease in response size in the lateral leads, and the second principal component corresponds to the change in response waveform at the lateral leads. There is little appreciable difference between the invariant description of the scalp topography elicited by FM and NFM.

To make this observation rigorous, we need to compare weights across subjects and conditions (FM *vs.* NFM) in a manner independent of the choice of basis vectors for the representations, eq. (2), for each subject. Further details on this procedure may be found in Victor and Conte (1991). Briefly, for any pair of subjects s_1 and s_2 and conditions c_1 and c_2, we form the dot product of weights from any pair of principal components, which we denote $D_{h_1,h_2}(s_1c_1, s_2c_2)$:

$$D_{h_1,h_2}(s_1c_1,s_2c_2) = \sum_d w_{h_1,s_1c_1}(d) \cdot w_{h_2,s_2c_2}(d). \qquad (6)$$

The dot product $D_{h_1,h_2}(s_1c_1, s_2c_2)$ indicates the correspondence of the weights $w_{h_1,s_1c_1}(d)$ for principal component h_1 from subject s_1 and condition c_1 with the weights $w_{h_2,s_2c_2}(d)$ for principal component h_2 from subject s_2 and condition c_2. The dot product has a value of 1 if the weights coincide, and a value of 0 if they are orthogonal. Thus, if only the first principal component in the expansion, eq. (2), is considered, then the similarity of scalp topographies across a pair of subjects and conditions is expressed by the cross invariant $X^{[1]}$:

$$X^{[1]}(s_1c_1,s_2c_2) = D_{1,1}(s_1c_1,s_2c_2). \qquad (7)$$

For two or more principal components, determinants are needed to eliminate the intrinsic ambiguity of the expansion, eq. (2). Cross invariants for expansions, eq. (2), consisting of two or three principal components are given by

$$X^{[2]}(s_1c_1,s_2c_2) = \det \begin{vmatrix} D_{1,1}(s_1c_1,s_2c_2) & D_{1,2}(s_1c_1,s_2c_2) \\ D_{2,1}(s_1c_1,s_2c_2) & D_{2,2}(s_1c_1,s_2c_2) \end{vmatrix} \qquad (8)$$

and

$$X^{[3]}(s_1c_1,s_2c_2) = \det \begin{vmatrix} D_{1,1}(s_1c_1,s_2c_2) & D_{1,2}(s_1c_1,s_2c_2) & D_{1,3}(s_1c_1,s_2c_2) \\ D_{2,1}(s_1c_1,s_2c_2) & D_{2,2}(s_1c_1,s_2c_2) & D_{2,3}(s_1c_1,s_2c_2) \\ D_{3,1}(s_1c_1,s_2c_2) & D_{3,2}(s_1c_1,s_2c_2) & D_{3,3}(s_1c_1,s_2c_2) \end{vmatrix} \qquad (9)$$

The values of the cross invariants do not depend on the choice of basis vectors within the H-dimensional subspace in the representation, eq. (2). Geometrically, the cross invariant $X^{[H]}$ is the cosine of the dihedral angle between the response subspaces for subject/condition s_1c_1 and for subject/condition s_2c_2. When the cross invariant is equal to 1, these subspaces coincide. That is, the scalp topographies correspond to an equivalent set of spatial weights (although the temporal waveforms of the responses may be different). When the cross invariant is equal to 0, the subspaces are perpendicular. In this case, the scalp topographies are uncorrelated. For example, if the first and third principal components for two conditions coincide, but the second principal components for two conditions are orthogonal, then $X^{[1]} = 1$ but $X^{[2]} = 0$ and $X^{[3]} = 0$. That is, the cross invariant $X^{[H]}$ provides a way to compare the first H principal components as a set, and not just the Hth component.

Table 1. *Principal components analysis of scalp topography of FM and NFM responses*

Subject	Number of components					
	1	2	3	1	2	3
	A. Cumulative fraction of explained variance					
	FM			NFM		
JV	0.957	0.988	0.998	0.973	0.990	0.998
MC	0.842	0.979	0.990	0.855	0.972	0.993
RB	0.976	0.983	0.997	0.975	0.988	0.995
	B. Cross-invariant analysis: FM *vs.* NFM					
	Cross-invariant value			Resampled P value		
JV	0.999	0.976	0.916	0.62	0.65	0.19
MC	0.983	0.972	0.963	0.21	0.06	0.41
RB	0.997	0.781	0.551	0.23	0.13	0.23
	C. Cross invariants for intersubject comparisons[a]					
	FM			NFM		
JV *vs.* MC	0.996**	0.914**	0.807	0.995**	0.962**	0.901*
JV *vs.* RB	0.982**	0.860**	0.713	0.996**	0.950**	0.541
MC *vs.* RB	0.995**	0.790**	0.704	0.994**	0.948**	0.541

[a]In Part C, significance of cross-invariant values is indicated by "*" ($P = 0.05$) and "**" ($P = 0.01$).

Part B of Table 1 displays the cross invariants $X^{[1]}$, $X^{[2]}$, and $X^{[3]}$ for comparisons of responses to FM and NFM within each subject. For each subject, $X^{[1]}$ is at least 0.98, $X^{[2]}$ is at least 0.78, and $X^{[3]}$ is at least 0.55. To determine whether any of these cross invariants demonstrate a significant difference in scalp topography between FM and NFM responses, we use a resampling scheme. The original data set for each subject consisted of four responses to FM and four responses to NFM. From these eight responses, we create artificial data sets by randomly assigning each of them to "condition A" or "condition B." Resampled cross invariants are then calculated from a set of randomly segregated responses. If cross invariants calculated from the resampled data are usually larger than cross invariants calculated from the actual data, we conclude that there are significant differences between the topography of FM and NFM responses. Conversely, if cross invariants calculated from the resampled data are similar to those calculated from the actual data, then variations within each set of responses (FM and NFM) are comparable to differences between FM and NFM responses. The fraction of 100 resampled cross invariants which is smaller than the actual cross invariant is the p-value listed

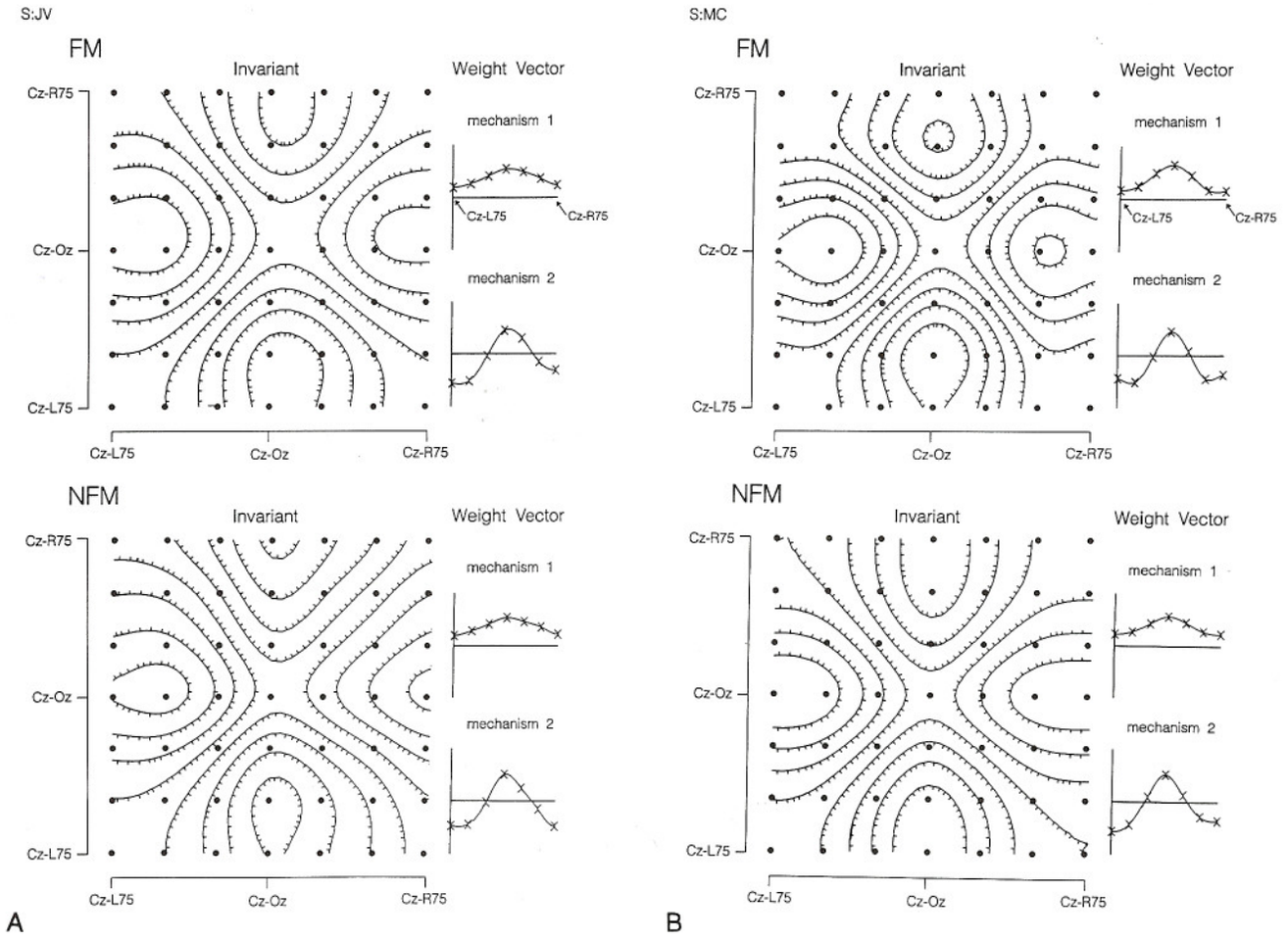

Fig. 9. Weight vectors and invariants $I^{[2]}(d,d')$ of the first two principal components derived from the data of Fig. 8 at contrast 0.4. The size of the invariant $I^{[2]}(d,d')$ represents the extent to which the first two principal components produce distinct dynamics at derivations d and d'. Each contour line represents a value of 0.1; tickmarks point downhill. The diagonal symmetry of the contour map is implied by the relationship $I^{[2]}(d,d') = -I^{[2]}(d',d)$. Subjects: JV (A) and MC (B).

in the Part B of Table 1. By this test, no cross invariant showed a significant difference between FM and NFM at $P = 0.05$.

We conclude the analysis of scalp topography by examining the consistency of the first three principal components across subjects. Cross invariants were calculated across each pair of subjects, but within each condition (FM or NFM). To assess whether the first principal components were consistent, we compared the calculated cross invariant $X^{[1]}$ with the distribution of values that would be obtained from random weight vectors w_{1,s_1} and w_{1,s_2}. To assess the incremental consistency of the second principal component, we compared the calculated $X^{[2]}$ with the distribution of values that would be obtained from a pair of expansions, eq. (2), for which the first weight vectors w_{1,s_1} and w_{1,s_2} coincide, and the second weight vectors w_{2,s_1} and w_{2,s_2} are random. A similar procedure was used to assess the incremental consistency of the third principal component. Further details may be found in Victor and Conte (1991). Critical values at $P = 0.05$ for $X^{[1]}$, $X^{[2]}$, and $X^{[3]}$ are 0.616, 0.667, and 0.723, respectively, and critical values at $P = 0.01$ are 0.720, 0.780, and 0.843, respectively. As seen in the Part C of Table 1, in all comparisons, the first two principal components were consistent across subjects at $P = 0.01$. The three-component expansion was only significant at $P = 0.05$ in one case (NFM, JV *vs.* MC). Thus, from the analysis of consistency across subjects, retention of the second term in the expansion, eq. (2), is fully justified, but retention of the third term is at best marginally justified.

In sum, we found (1) that a principal-components expansion in two or three components represented at least 99% of the power (and thus accurately represented scalp topography); (2) there were no significant differences between the scalp topography of FM and NFM responses as represented by one-, two-, or three-component expansions; and (3) the first two principal components were consistent across subjects.

Directionally selective responses?

If directionally selective responses were present in the VEP, they might be expected to be found only in the laterally placed leads, since bilateral symmetry would cause cancellation of such signals at the midline. The change in response waveform with electrode location might be because responses recorded more laterally contained directionally selective components. We therefore subtracted signals recorded during rightward drift from

those recorded during leftward drift (see Methods). As seen in Fig. 10, no such directionally selective components were apparent under these recording conditions for FM or NFM. We used the T^2_{circ} statistic to make this observation more precise. The first four Fourier components of the left-right difference waveform were examined from three subjects, two conditions (FM and NFM), and the seven electrode sites. Of the 4 Components × 3 Subjects × 2 Conditions × 7 Sites = 168 Fourier Components, only six were significant at $P = 0.05$; chance alone would typically result in eight "significant" values at this confidence level.

Dependence on textural cues?

We previously showed (Victor & Conte, 1990) that the detection of motion in NFM stimuli depended critically on what kind of cues were available: contours signaled by flicker or granularity provided a strong percept of motion, while contours signaled by higher-order elements of form did not generate a strong movement percept. We therefore asked how these cues influenced the VEP.

The stimuli consisted of the first set of variations described in Methods and illustrated in Fig. 2. These NFM stimuli consist of alternate bands of regions filled in by the even texture (Julesz et al., 1978), a random coloring of individual pixels, or a random coloring on a coarser scale. As seen in the first column of Fig. 11, all three NFM stimuli (*E/R*-band, *C/R*-band, and *R/R*-band) elicited a large motion-onset VEP, similar to those elicited by the standard NFM stimulus *R/R*-drift*. In all of these stimuli, onset of motion necessarily coincided with onset of flicker, since the motion of the strip boundaries necessarily required recoloring of individual pixels. To remove this cue, we modified the stimulus so that all pixels were subject to recoloring at each drift step (but the recolored checks continued to conform to the rule of the strip in which they were situated).

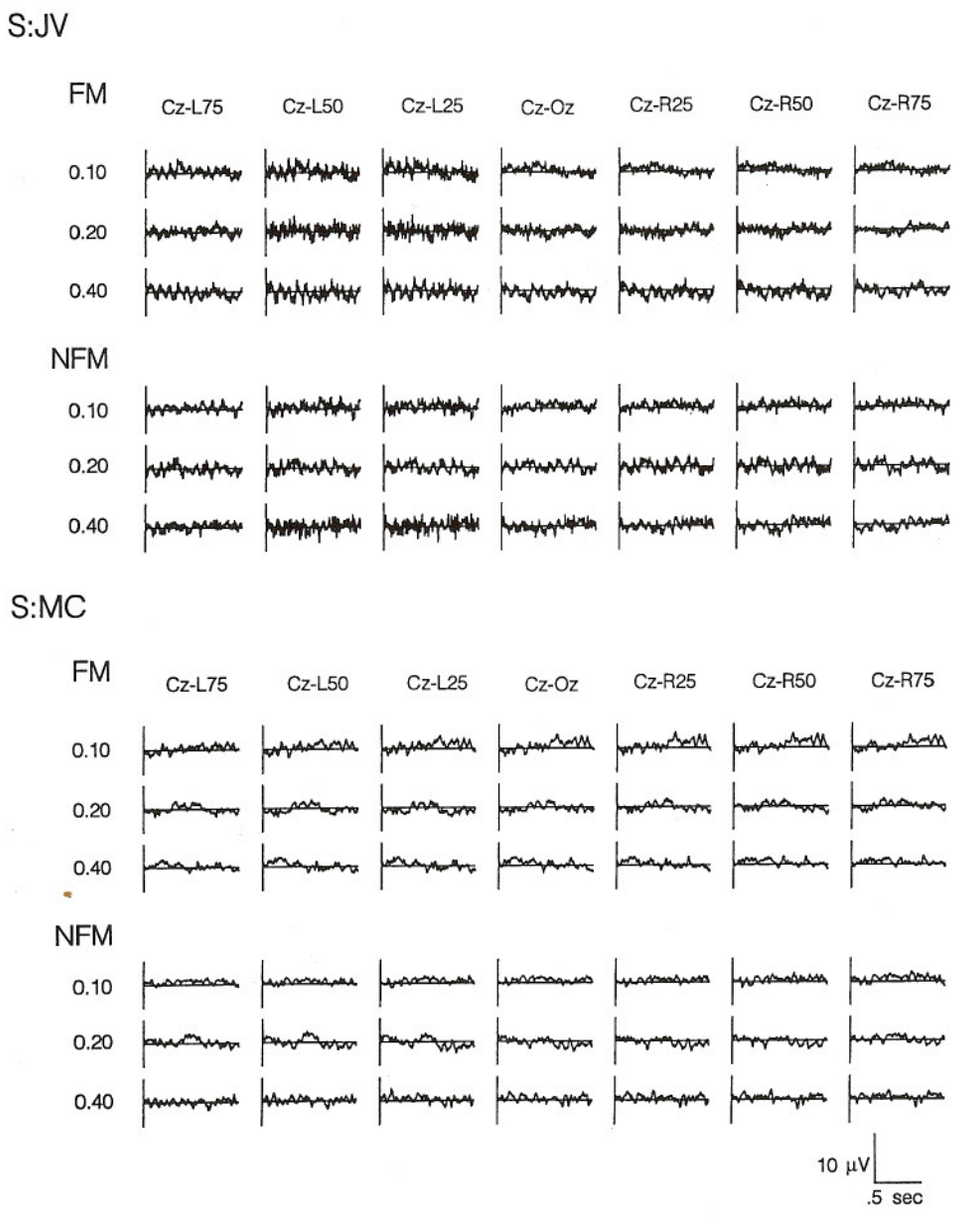

Fig. 10. Differences of responses to leftward and rightward drift elicited by FM (*R/R*-drift) and NFM (*R/R*-drift*) stimuli. Contrast = 0.4. Onset of motion occurs at the beginning of each histogram; cessation of motion occurs in the middle of each histogram. Each histogram represents a difference of the average response to at least 128 epochs of leftward drift and 128 epochs of rightward drift. Subjects: JV and MC.

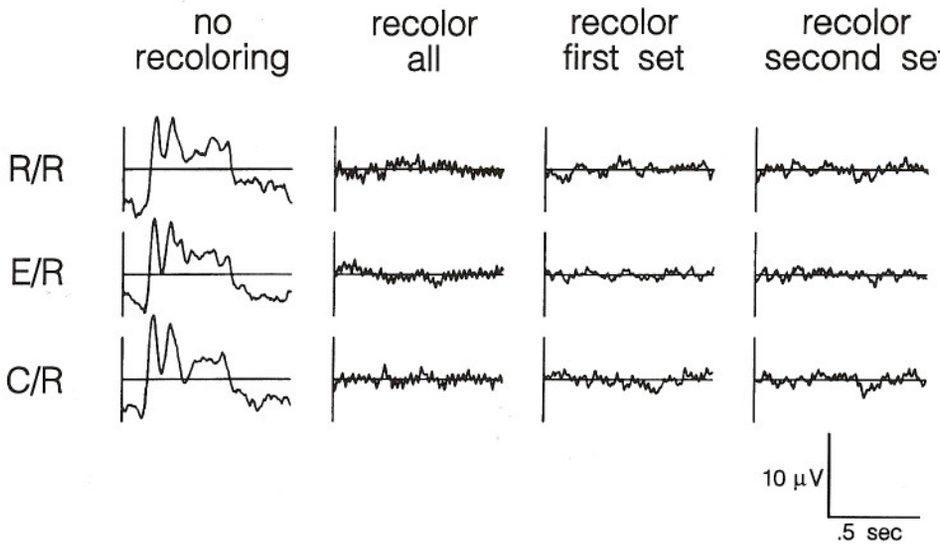

Fig. 11. VEPs elicited by NFM stimuli in which the texture strips were recolored from frame to frame. First column: no recoloring (*R/R*-band, *E/R*-band, *C/R*-band). Second column: all strips recolored (*R*/R**-band, *E*/R**-band, *C*/R**-band). Third column: first set of strips recolored (*R*/R*-band, *E*/R*-band, *C*/R*-band). Fourth column: second set of strips recolored (*R/R**-band, *E/R**-band, *C/R**-band). Each histogram represents an average of at least 256 epochs of drift, with leftward and rightward drift superimposed. Derivation: C_z-O_z. Contrast: 0.4. Subject: MC.

For example, in the E^*/R^*-band stimulus, each even strip was recolored with a new even texture and each random strip was recolored randomly every 29.6 ms. Since the recoloring persisted during the period in which the strips no longer drifted, flicker no longer provided a cue to motion. Thus, in the stimulus E^*/R^*-drift, the only cue to motion is a higher-order form cue (stimulus *H* of Victor & Conte, 1990). The maneuver of recoloring each strip on each frame also produces the stimulus R^*/R^*-drift, which has no cues to motion, and the stimulus C^*/R^*-drift, which has only a granularity cue to motion (stimulus *G* of Victor & Conte, 1990). As seen in the second column of Fig. 11, none of these stimuli produced an appreciable motion-onset VEP. This does *not* correspond to the psychophysical findings (Victor & Conte, 1990). Psychophysically, a strong motion percept is present when the flicker cue is removed provided that a granularity cue remains. The granularity cue remains in the stimulus C^*/R^*-drift, but this stimulus does not produce a motion-onset VEP.

The dissociation between psychophysics and the VEP was further explored by recoloring only one set of vertical strips. In these stimuli, flicker provided a cue to motion, since alternate vertical strips were recolored every frame but the intervening strips were not. Psychophysically, the percept of motion for these stimuli was comparable to that for the unflickered stimuli (Table 2). Nevertheless, none of these stimuli elicited a significant VEP.

Similar results were obtained in the other two subjects. As above, we used the sum of the amplitudes of the first four Fourier components to quantify response size, and the T^2_{circ} statistic to determine the statistical significance of individual responses. For the midline lead (which contained the largest responses), responses to the unflickered stimuli had a mean of 6.2 μV (MC), 6.7 μV (JV), and 3.3 μV (RB). Of the 4 Components × 3 Subjects × 3 Conditions = 36 Components, 30 were significantly different from zero at $P = 0.05$. (The Fourier coefficients that were zero were typically the third- or fourth-order coefficients.) In contrast, when either or both of the sets of bars were recolored each frame, responses had a mean of less than 0.9 μV in all subjects. Analysis by the T^2_{circ} statistic revealed that, at best, minimal responses to these stimuli may have been present. Of the 4 Components × 3 Subjects × 9 Conditions = 108 Components, nine were significantly different from zero at $P = 0.05$. By chance alone, nine or more coefficients out of 108 would be significant approximately 9% of the time. Thus, while psychophysical responses to motion persist after manipulations which eliminate onset of flicker as a cue to onset of motion, this manipulation essentially eliminates the VEP response to NFM.

We also examined whether the responses to the three stimuli (*R/R*-band, *E/R*-band, *C/R*-band) which generated a motion-onset VEP were significantly different. At the midline lead, of the 4 Components × 3 Subjects × 3 Conditions pairs = 36 Component pairs, only three were significantly different at $P = 0.05$.

Principal components analysis was used to compare scalp topographies across these conditions (Table 3). As in the responses to the standard stimulus, two principal components typically fit at least 98% of the response power. Within-subject analysis of the cross invariants revealed only one of 27 possible comparisons to be significantly different at $P = 0.05$. Thus, there is no apparent difference between the dynamics or the scalp topography of responses to NFM stimuli which contain granularity or higher-order form cues, and those which do not.

Table 2. *Fraction of correct detections of direction of motion after brief masked presentations of stimuli R/R-band and R/R*-band*

	Subject					
	JV		MC		RB	
Stimulus	*R/R*	*R/R**	*R/R*	*R/R**	*R/R*	*R/R**
Presentation time (ms)						
59	0.45	0.55	0.45	0.53	0.45	0.38
89	0.68	0.75	0.72	0.67	0.63	0.53
118	0.95	0.95	0.98	0.82	0.93	0.63
237	1.00	1.00	1.00	1.00	1.00	0.97

Table 3. *Principal components analysis of scalp topography of responses to NFM stimuli E/R-band, R/R-band, and C/R-band*

A. Cumulative fraction of explained variance

	Number of components								
Subject	1	2	3	1	2	3	1	2	3
	E/R-band			*R/R*-band			*C/R*-band		
JV	0.958	0.988	0.997	0.969	0.990	0.996	0.954	0.988	0.995
MC	0.928	0.985	0.993	0.913	0.985	0.993	0.883	0.983	0.992
RB	0.924	0.965	0.992	0.939	0.979	0.997	0.889	0.946	0.973

B. Cross-invariant analysis

	Number of components					
	1	2	3	1	2	3
	Cross-invariant value			Resampled *P* value		
	R/R-drift* *vs.* *R/R*-fixed*					
JV	0.994	0.856	0.978	0.00	0.32	0.79
MC	0.999	0.996	0.967	0.31	0.70	0.01
RB	0.998	0.913	0.923	0.09	0.45	0.22
	E/R-band *vs.* *C/R*-band					
JV	0.999	0.792	0.946	0.57	0.00	0.20
MC	0.999	0.989	0.936	0.39	0.19	0.49
RB	0.997	0.817	0.880	0.18	0.25	0.29
	R/R-band *vs.* *C/R*-band					
JV	0.999	0.979	0.977	0.30	0.66	0.62
MC	0.997	0.991	0.966	0.09	0.29	0.75
RB	0.996	0.695	0.819	0.16	0.06	0.10

Flicker-defined motion vs. flicker-defined contour

The stimuli (standard stimulus *R/R*-drift* and those used in the first column of Fig. 11) which generate a motion-onset VEP all have the onset of a flicker-defined contour at the same time as the motion begins. The stimuli in which one or more strips are recolored each frame (last three columns of Fig. 11) and do not generate a motion-onset VEP contain a flicker-defined contour throughout the entire stimulus cycle. Thus, one interpretation for the elimination of the motion-onset VEP under these conditions is that the "motion-onset" VEP might be driven by the onset of flicker-defined contours, rather than non-Fourier motion. That is, the response to a moving flickering contour might be identical to the response to a stationary flickering contour.

To test this hypothesis, we used the second set of stimulus variations described in Methods and illustrated in Fig. 3. The stimulus *R/R*-fixed* (Fig. 3A) differs from the standard NFM stimulus *R/R*-drift* in that the flicker-defined contour does not undergo apparent motion; rather, it appears at the beginning of each averaging period and disappears halfway through each averaging period. Necessarily, the contours in the stimuli *R/R*-drift* changed in position, while the contours in the stimulus *R/R*-fixed* did not. The contours in each stimulus were separated by approximately 1 deg. Since we wanted to look for dependence on motion and not spatial position, we therefore presented these stimuli in each of four spatial phases one-quarter cycle (16 min) apart. Fig. 12 shows responses to these stimuli at each of four spatial phases, as well as the grand average across spatial phases, for subject RB. A mild dependence on spatial phase is present—for example, the transient in the response to the stimulus *R/R*-fixed* is smallest in spatial phase 2. However, in all spatial phases, the initial transient is larger for the response to *R/R*-drift* than for the response to *R/R*-fixed*. This implies that the difference in responses to these two stimuli was related to the attribute of motion, and not simply due to the different locations in the visual field that were stimulated.

We also compared the responses to smooth NFM (*R/R*-drift*) with responses elicited by a flicker-defined bar which moved randomly (*R/R*-jitter*, Fig. 3B). Responses to *R/R*-jitter* show a spatial-phase dependence (third row of Fig. 12), but in all spatial phases, the initial transient in the response to *R/R*-drift* is larger than the initial transient in the response to *R/R*-jitter*. The transient parts of the *R/R*-fixed* response and the *R/R*-jitter* response were similar, but the response to *R/R*-jitter* remained above baseline for a longer time than the response to *R/R*-fixed* (last column of Fig. 12).

Similar results were found in the other subjects. We summarize the dependence of the responses on stimulus type and spatial phase *via* statistical comparisons of their first four Fourier

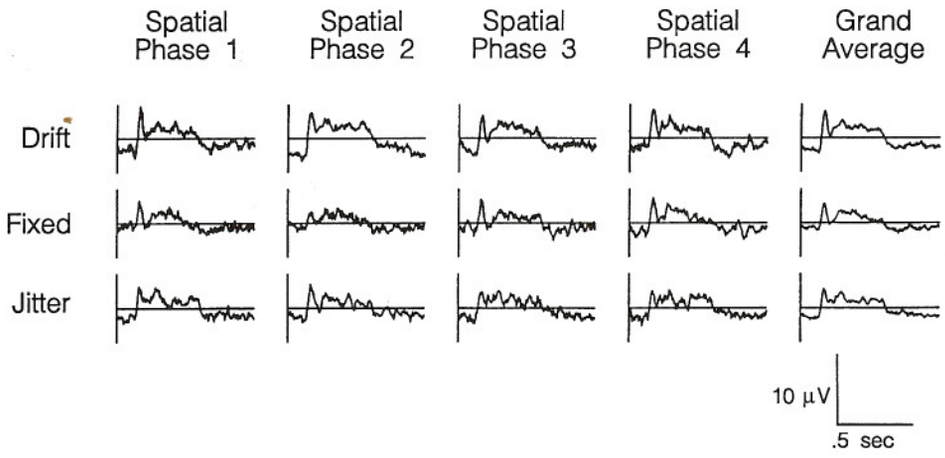

Fig. 12. VEPs elicited by the stimuli *R/R*-drift*, *R/R*-fixed*, and *R/R*-jitter*. Onset of motion or flicker occurs at the beginning of each histogram; cessation occurs in the middle of each histogram. Each histogram in the first four columns represents an average of responses to 192 epochs in which the stimulus was presented in a single spatial phase. The rightmost column is an average over all spatial phases. In spatial phase 1, the contour appears in the midline. In each successive spatial phase, the position of initial contour appearance is shifted by one-quarter cycle (16 min) left. Derivation: C_z-O_z. Subject: RB.

components. We use the T^2_{circ} statistic and a significance level of $P = 0.05$. A significant dependence on spatial phase was seen in 173 comparisons out of 4 Components × 3 Subjects × 3 Stimulus types × 6 Phase pairs × 7 Sites = 1512, or slightly more than 11%. There was a larger difference across conditions (drift *vs.* fixed *vs.* jitter). Restricting comparisons to within individual spatial phases, there was a significant dependence on stimulus type in 193 comparisons out of 4 Components × 3 Subjects × 3 Stimulus-type Pairs × 4 Phases × 7 Sites = 1008, or 19%. We also compared responses to the three stimulus types after pooling across spatial phases. Pooling across spatial phases provides added power in that more repetitions are combined for the estimate of each Fourier component. Yet, since this manipulation considers spatial-phase dependence to be a random source of intertrial variability, it is a conservative test of the dependence of dynamics on stimulus type. There was a significant dependence on stimulus type in 112 comparisons out of 4 Components × 3 Subjects × 3 Stimulus-type pairs × 7 Sites = 252, or 44%. Differences were approximately equally divided among all three pairwise comparisons (44 for drift *vs.* fixed, 30 for drift *vs.* jitter, 38 for fixed *vs.* jitter) and subjects (36 for JV, 37 for MC, 39 for RB).

We examined the dependence of scalp topography on stimulus type through principal components analysis and cross invariants. For the reasons mentioned above, we pool responses across spatial phase for this analysis. As seen in Table 4, two principal components suffice to account for 97–98% of response power, and three principal components typically account for greater than 99% of response power. Resampling statistics for the cross invariants demonstrate that seven of the possible 27 comparisons (3 Subjects × 3 Stimulus-type pairs × 3 Cross invariants) are significant at $P = 0.05$, and four of these are significant at $P = 0.01$. All of the significant differences were in subjects JV and MC. This indicates that in these subjects, there is also a dependence of scalp topography on stimulus type, in addition to the dependence of response dynamics on stimulus type.

Table 4. *Principal components analysis of scalp topography of responses to the NFM stimulus R/R-drift* and related stimuli R/R-fixed* and R/R-jitter* without coherent motion*

A. Cumulative fraction of explained variance

	Number of components								
Subject	1	2	3	1	2	3	1	2	3
	R/R-drift*			*R/R*-fixed*			*R/R*-jitter*		
JV	0.948	0.979	0.991	0.973	0.988	0.995	0.925	0.977	0.991
MC	0.906	0.978	0.994	0.920	0.976	0.994	0.884	0.956	0.985
RB	0.951	0.981	0.995	0.918	0.967	0.991	0.951	0.981	0.995

B. Cross-invariant analysis

	Number of components					
	1	2	3	1	2	3
	Cross-invariant value			Resampled P value		
	R/R-drift* *vs.* *R/R*-fixed*					
JV	0.994	0.856	0.978	0.00	0.32	0.79
MC	0.999	0.996	0.967	0.31	0.70	0.01
RB	0.998	0.913	0.923	0.09	0.45	0.22
	R/R-drift* *vs.* *R/R*-jitter*					
JV	0.998	0.932	0.795	0.27	0.35	0.24
MC	0.999	0.936	0.986	0.43	0.00	0.34
RB	0.998	0.724	0.990	0.13	0.18	0.71
	R/R-fixed* *vs.* *E/R*-jitter*					
JV	0.995	0.753	0.765	0.01	0.27	0.16
MC	0.997	0.958	0.965	0.04	0.04	0.00
RB	0.999	0.920	0.910	0.52	0.45	0.14

Discussion

Summary of results

We presented the results of three experiments.

1. We compared VEPs elicited by FM and NFM stimuli matched for spatial and temporal characteristics (Fig. 1). Over a wide range of velocities, VEP amplitude and waveforms elicited were indistinguishable (Figs. 4–6), and there were no significant differences in the VEP scalp topography as measured at a transverse chain of electrodes (Figs. 8–10; Table 1).
2. We examined VEPs elicited by NFM stimuli (Fig. 2) in which textural cues such as granularity and higher-order form elements were added, and in which the onset of flicker as a cue to motion was eliminated. We found that additional textural cues did not influence the response dynamics or scalp topography (Table 3), and that significant VEPs were present only when the onset of motion coincided with the onset of flicker (Fig. 11).
3. We compared responses to NFM stimuli and related stimuli without coherent motion but which shared the spatial and temporal characteristics of the NFM stimuli (Fig. 3). There were significant differences in the dynamics in all subjects (see Fig. 12 and related text) and scalp topography of the responses to these stimuli in two subjects of three (Table 4). This demonstrated that the motion-onset responses seen for NFM stimuli were indeed distinct from responses to the onset of a fixed flicker-defined contour, or, to the onset of a flicker-defined contour moving randomly.

Caveats

The motion-related VEPs presented here resemble those recorded subdurally from macaque V1 (van Dijk et al., 1986; van Dijk & Spekreijse, 1989) in overall waveform, in the relative insensitivity to velocity, and in the more transient and larger response at motion onset compared with that at motion cessation. Thus, it is reasonable to consider the VEP to represent activity of an early stage of visual processing. However, since the scalp recorded VEP fundamentally is a composite of activity of many neural populations, assignment of the scalp VEP to any particular stage of processing must be viewed with great caution. Nevertheless, it is a measure which provides a window on dynamics difficult to achieve by standard psychophysical techniques, as well as crude information concerning brain localization. The

finding that VEP dynamics and topography of responses to FM and NFM stimuli are similar over a wide range of velocity and contrast thus complements psychophysical evidence concerning the similarity of Weber fractions for velocity discrimination of FM and NFM stimuli (Turano & Pantle, 1989; Turano, 1991).

We recognize that failure to find a difference in scalp topography does not imply that two sets of generators are identical. However, we note that the method we have used (cross invariants and resampling statistics) is sensitive enough to demonstrate interocular differences in the cortical representation of the visual hemifield of ipsilateral and contralateral hemiretinae (Victor et al., 1991), and also succeeds in demonstrating a difference between responses to coherently moving, stationary, and randomly moving contours (Table 4).

Physiological implications

The basic motivation for this study was to investigate the relationship of pathways which processed NFM to pathways which processed FM. One view is that FM, which is well accounted for by Reichardt-type cross correlators, represents processing of motion over short intervals of space and time (the short-range system of Braddick, 1974). In this view, NFM represents a separate system which operates over longer ranges, and perhaps at a later stage of processing. However, our finding of a striking identity between the dependence on velocity and contrast of the responses to both kinds of motion makes this view unattractive, and strongly implies that these two kinds of motion are processed by the same pathway.

Chubb and Sperling (1989) pointed out that NFM can be extracted by a Reichardt-like cross correlator, provided that the cross correlator is preceded by a local nonlinearlity and additional temporal filtering. A parsimonious account of our results is that the local nonlinearity is an asymmetric one. In this way, it will transmit signals linearly related to the stimulus along with the nonlinear component. The same cross-correlator stage will then suffice to extract FM (from the linear signals) and NFM (from the nonlinear signals). Thus, spatial tuning and dynamics of responses to FM and NFM stimuli are necessarily linked. This view is consistent with the findings of Albright (1992), who found that individual neurons in area MT of macaque cortex which were sensitive to FM stimuli were typically comparably sensitive to NFM stimuli. As we will see below, this has important consequences for computational modeling of the extraction of motion signals.

Motion onset vs. contour onset

A second motivation of these studies was to determine to what extent VEPs elicited by the onset of motion were driven by motion *per se*, and not merely by contrast changes. NFM stimuli provided a new way to examine differences between responses driven by motion and responses driven by contours, since with these stimuli, it is possible to dissociate appearances of contours from the onset of motion. As shown in Fig. 11 (columns 2, 3, and 4), a NFM stimulus fails to elicit a detectable VEP when the onset of motion is unaccompanied by the appearance of a contour. This is true even when the onset of motion in the NFM stimulus is readily detectable (columns 3 and 4 of Fig. 11). This finding supports the conclusion of Spekreijse et al. (1985) that there was little to distinguish the motion-onset response from the pattern-reversal response. Indeed, these investigators considered the standard contrast-reversal VEP response to be adequately described in terms of motion onset and motion offset.

Another way to address the motion specificity of the response is to compare the response to onset of NFM to the response to the onset of a stationary contour defined by random flicker. That is, we compare the response to onset of motion of a non-Fourier contour with the response to appearance of a non-Fourier contour. As seen in Fig. 12, appearance of a contour defined by random flicker evokes a robust VEP, whether the contour is drifting coherently (R/R-drift*), stationary (R/R-fixed*), or moving randomly (R/R-jitter*). Though these responses are qualitatively similar, there were statistically significant differences in their dynamics and scalp distribution. The early portion of the response to stimuli with coherent motion was larger than that elicited by stimuli without coherent motion (Fig. 12). These findings, especially when taken in combination with responses obtained during tracking of the moving contour (Fig. 7), indicate that at least some of the response is driven by motion *per se*, and not merely by spatiotemporal contrast changes. Furthermore, they demonstrate at least subtle differences in the response to moving contours and responses to static contours with identical spatial and temporal characteristics. One possible explanation for this finding is that the differences in scalp topography are a result of relative insensitivity of the foveal representation to motion onset (van Dijk et al., 1986). A second possibility is that these differences represent specific involvement of extrastriate regions—for example, the human analogue of MT—in the processing of the moving stimuli.

Computational implications

We now consider in more detail the notion that a standard motion model, preceded by an appropriate nonlinearity, accounts for our data. This will lead to qualitative conclusions concerning the form of the nonlinearities themselves. In this analysis, we replace the cross correlator of the Reichardt (1961) model by calculation of motion energy in opponent directions (Adelson & Bergen, 1985). This is in keeping with recent physiological evidence (Emerson et al., 1987, 1992) that directionally selective complex cells extract motion energy in a single direction, and do not behave as bidirectional cross correlators *per se*.

For analytical purposes, the model for motion extraction is as follows:

i. The first stage of the model is a linear spatiotemporal filter, which is bandpass both spatially and temporally. This model stage represents retinal and other stages of early visual processing and is not specific to motion. Since this filter is considered to have a spatiotemporal impulse response which is not oriented in space-time, it cannot contribute to directionally selective responses.

ii. The second stage of the model is a local nonlinearity, whose input-output relation is

$$u = N_{\text{local}}(s). \tag{10}$$

iii. The output of the local nonlinearity is filtered by a directionally selective linear filter. That is, this linear filter has

a spatiotemporal impulse response which consists of oblique bands in space-time.

iv. Finally, the output of the directionally selective linear filter is subject to a second nonlinearity,

$$u = N_{\text{energy}}(s), \tag{11}$$

to generate a motion-energy signal. $N_{\text{energy}}(s)$ should always be nonnegative.

It is generally assumed (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Emerson et al., 1992) that the second nonlinearity N_{energy} is a squaring operation, so we consider this possibility first.

The local nonlinearity $N_{\text{local}}(s)$ can be decomposed into an even-symmetric component $N_{\text{even}}(s)$ [satisfying $N_{\text{even}}(s) = N_{\text{even}}(-s)$] and an odd-symmetric component $N_{\text{odd}}(s)$ [satisfying $N_{\text{odd}}(s) = -N_{\text{odd}}(-s)$]. That is, $N_{\text{even}}(s)$ is independent of the polarity of the signal s (and therefore of the stimulus itself), while $N_{\text{odd}}(s)$ inverts its signature with reversal of the polarity of stimulus contrast. The decomposition is

$$N_{\text{local}}(s) = N_{\text{even}}(s) + N_{\text{odd}}(s), \tag{12}$$

where

$$N_{\text{even}}(s) = \tfrac{1}{2}[N_{\text{local}}(s) + N_{\text{local}}(-s)]$$

and

$$N_{\text{odd}}(s) = \tfrac{1}{2}[N_{\text{local}}(s) - N_{\text{local}}(-s)].$$

This decomposition facilitates an analysis of the action of the second nonlinearity N_{energy} on the output of the local nonlinearity. Since the second nonlinearity is assumed to be quadratic, only pairwise products of its inputs need to be considered.

Consider a pair s_1 and s_2 of inputs to the local nonlinearities at two points (x_1, t_1) and (x_2, t_2) of space-time. That is, s_1 and s_2 represent the signals resulting from transformation of the stimulus by the linear spatiotemporal filter (i). The pairwise product $N_{\text{local}}(s_1)N_{\text{local}}(s_2)$ may be decomposed according to eq. (12) as follows:

$$\begin{aligned} &N_{\text{local}}(s_1)N_{\text{local}}(s_2) \\ &\quad = N_{\text{even}}(s_1)N_{\text{even}}(s_2) + N_{\text{odd}}(s_1)N_{\text{odd}}(s_2) \\ &\qquad + N_{\text{even}}(s_1)N_{\text{odd}}(s_2) + N_{\text{odd}}(s_1)N_{\text{even}}(s_2). \end{aligned} \tag{13}$$

We now consider the contribution of each of these terms to a motion-energy signal, averaged over the stimulus area, for the stimuli used in these studies. The two cross terms $N_{\text{even}}(s_1) \times N_{\text{odd}}(s_2)$ and $N_{\text{odd}}(s_1)N_{\text{even}}(s_2)$ each have the property that if the stimulus contrast is changed in polarity, then the contribution of the term is inverted in sign. The stimuli used in these studies are left invariant by a change in overall polarity of contrast—they are all constructed based on an initial random assignment of checks to positive and negative contrasts. Since these assignments are uncorrelated over space, a spatial average of either of the cross terms must tend to zero.

Thus, for a quadratic second nonlinearity N_{energy}, the only contributions to direction selectivity are the result of the terms $N_{\text{even}}(s_1)N_{\text{even}}(s_2)$ and $N_{\text{odd}}(s_1)N_{\text{odd}}(s_2)$. For FM stimuli, both terms may contribute. The contribution of $N_{\text{odd}}(s_1)N_{\text{odd}}(s_2)$ may be revealed by the choice $N_{\text{local}}(s) = s$. For this choice, $N_{\text{even}} = 0$ and the standard motion-energy model (without any initial nonlinearity) is recovered. The contribution of $N_{\text{even}}(s_1)N_{\text{even}}(s_2)$ may be revealed by the choice $N_{\text{local}}(s) = |s|$. This choice (for which $N_{\text{odd}} = 0$) removes the polarity of the stimulus contrast or contrast change prior to standard motion analysis. For the FM stimuli used in these studies, direction of motion remains unambiguous following this maneuver, even though the contributions of the terms $N_{\text{odd}}(s_1)N_{\text{odd}}(s_2)$ are necessarily zero. This implies that $N_{\text{even}}(s_1)N_{\text{even}}(s_2)$ suffices to generate an unambiguous directionally selective signal.

The previous example demonstrates that the interactions $N_{\text{even}}(s_1)N_{\text{even}}(s_2)$ suffice to generate a directionally selective signal not only for FM stimuli, but also for NFM stimuli. The difference between the FM stimulus R/R-drift and the NFM stimulus R/R-drift* is that in the NFM stimulus, the contrast polarity of the moving check is randomized, while in the FM stimulus, it is fixed. That is, if contrast polarity is ignored, these two stimuli share the same spatiotemporal correlation structure. As a consequence, the interactions $N_{\text{even}}(s_1)N_{\text{even}}(s_2)$ provide the *same* contribution from FM and NFM stimuli.

The situation is quite different for the interactions $N_{\text{odd}}(s_1) \times N_{\text{odd}}(s_2)$. For the NFM stimuli, a sign-preserving initial nonlinearity (such as N_{odd}) results in a spatiotemporal pattern of activity with *no* second-order correlation along diagonal directions. This is because in the NFM stimulus, the moving checks vary randomly in polarity. As a consequence, the contribution of $N_{\text{odd}}(s_1)N_{\text{odd}}(s_2)$ must be zero.

To sum up, we have shown that under the assumption of a quadratic nonlinearity for motion-energy extraction N_{energy}, directionally selective signals generated by FM stimuli arise from terms $N_{\text{odd}}(s_1)N_{\text{odd}}(s_2)$ and $N_{\text{even}}(s_1)N_{\text{even}}(s_2)$. For the corresponding NFM stimuli, the contributions of the terms $N_{\text{odd}}(s_1)N_{\text{odd}}(s_2)$ are zero, while the contributions of the terms $N_{\text{even}}(s_1)N_{\text{even}}(s_2)$ match their contributions in the FM case. Thus, it is tempting to conclude that the initial nonlinearity N_{local} is purely even-symmetric—this would account for extraction of FM and NFM motion, and for the fact that the responses to these stimuli are nearly identical.

But there is a major problem with this view. As shown by Nakayama and Silverman (1985), direction selectivity for a luminance grating moving in discrete steps is optimal when the step size is $\frac{1}{4}$ of a cycle. For this apparent-motion stimulus, the positions of zero crossings on one frame become the positions of extrema (maxima and minima) on the next frame, and *vice versa*. The unambiguous cue to direction of stimulus motion is the polarity of each lobe of the grating. Preservation of this information requires a sign-preserving component in the local nonlinearity, such as would be generated by N_{odd}. A sign-ignoring nonlinear transformation of the stimulus (such as N_{even}) could only provide a sense of ambiguous stimulus motion, since it suppresses the contrast-polarity cue. Thus, we cannot accept the hypothesis that $N_{\text{odd}} = 0$; it would imply that the perceived direction of motion in the quarter cycle stimulus described above would be ambiguous.

There is, however, another possibility. The entire analysis above was predicated on extraction of motion energy by a quadratic nonlinearity N_{energy}. As we now show, replacement of

the quadratic nonlinearity by a nonlinearity which is less accelerating allows for the presence of both N_{odd} and N_{even}, *and* near equality of responses to FM and NFM motion.

To see how this works, we specify the four-stage model (i–iv) as follows. This model represents a caricature of a more realistic model (which would have more physiologic filtering stages), but serves to demonstrate the point without a heavy computational burden. The initial linear stage (i) is posited to produce a signal which is equal to the difference between its current (signed) contrast and its (signed) contrast on the previous frame. That is, for a stimulus of contrast c, possible values of the output of stage (i) and $+2c$ (stimulus contrast changes from $-c$ to $+c$), $-2c$ (stimulus contrast changes from $+c$ to $-c$) and 0 (stimulus contrast is unchanged). The local nonlinearity (ii) is posited to be a half-wave rectifier. That is, it produces an output which is equal to the input if the input is positive, and is otherwise zero. For this $N_{local}(s)$, the decomposition, eq. (12), is $N_{even}(s) = |s|/2$ and $N_{odd}(s) = s/2$. The directionally selective linear stage (iii) is posited to be a linear filter which sums the incoming values along a diagonal trajectory in space-time. We assume that this trajectory is parallel to the direction of motion (translation by one check in space for each frame) and is k checks long, with k a reasonably large number (e.g. 10). Finally, the motion-energy calculation is assumed to be full-wave rectification with a small threshold s_0:

$$N_{energy}(s) = \begin{cases} |s| - s_0, & |s| > s_0 \\ 0, & |s| \leq s_0. \end{cases} \tag{14}$$

The threshold s_0 will be specified below.

Except for the local nonlinearity (ii), this model is similar in overall structure to that proposed by Emerson et al. (1992) for directionally selective neurons in cat striate cortex. As in the Emerson model, the output represents a motion-energy signal in one direction of an opponent pair, and direction selectivity is conferred by a linear filter whose spatiotemporal impulse response is inseparable (i.e. elongated on a diagonal in space-time, and not equal to a product of a purely spatial and a purely temporal component). The main differences between our model and Emerson's are the presence of the local nonlinearity and the nature of the second nonlinearity. Although it was natural for Emerson et al. (1992) to propose a quadratic nonlinearity for motion-energy calculation on the basis of white-noise analysis, other nonlinearities would be equally consistent with their results, which were obtained at only one contrast level.

Consider the response of the proposed computational unit to the FM stimulus *R/R*-drift. We only need to consider the situation in which the second linear filter (iii) is coincident with the trajectory of a moving check or is one check behind the trajectory of a moving check; other placements of the filter yield a response of zero. Consider first a moving check of positive contrast, and the second linear filter positioned coincident with its trajectory. As the moving bright check passes across the random background, half of the time, it generates a signal $s = +2c$ (negative-contrast check becoming positive), and half of the time, it generates a signal $s = 0$ (positive-contrast check not changing). Following half-wave rectification by N_{local} [stage (ii)], this signal is unchanged. These individual responses are uncorrelated. Therefore, stage (iii) calculates a sum of k values, each of which is 0 or $+2c$, with equal probability. Since this sum is guaranteed to be positive, the rectification component of the nonlinearity N_{energy} at stage (iv) has no effect. The threshold s_0 is chosen to be comparable to $2c$. Provided k is large, then nearly all of the time, one or more of the k contributing values will be equal to $2c$. Under these conditions, the threshold of eq. (14) reduces to a subtraction. Thus, the average signal size generated by stage (iv) for this positioning of the linear filter (iii) and a moving bright check is $kc - s_0$.

Consider next repositioning the linear filter (iii) so that it lags one check behind the moving check. As the moving check passes across the static background, half of the time it generates a signal $s = -2c$ (positive-contrast check becoming negative) and half of the time, it generates a signal $s = 0$ (positive-contrast check not changing). Thus, the signal received by N_{local} is guaranteed to be nonpositive, and is truncated to zero. This position of the linear filter (iii) therefore generates no contribution to motion energy.

The situation of a moving negative-contrast check need not be separately considered. Trajectories lined up with a moving negative-contrast check produce the same response as trajectories one check behind a moving positive-contrast check, i.e. 0. Trajectories one check behind a moving negative-contrast check produce the same response as trajectories coincident with a moving positive-contrast check, i.e. $kc - s_0$. Finally, the net contribution to motion energy for the FM stimulus *R/R*-drift (summed over trajectory positions, averaged over checks of positive and negative contrast) is $kc - s_0$.

Now we examine the response of this model to the NFM stimulus *R/R*-drift*. As the moving flickering check passes across the random background, one-quarter of the time, it generates a signal $s = +2c$ (negative-contrast check becoming positive), one-quarter of the time, it generates a signal $s = -2c$, (positive-contrast check becoming negative), and half of the time, it generates a signal $s = 0$ (positive-contrast or negative-contrast check not changing). Following half-wave rectification by N_{local} [stage (ii)], there is a signal of $s = +2c$ one-quarter of the time, and $s = 0$ three-quarters of the time. Summation of k of these signals at stage (iii) produces a distribution of responses whose mean is $\frac{1}{4}(k)(2c)$, or $kc/2$. Since these signals can never be negative, the rectification component of N_{energy} at stage (iv) has no effect. Since (for large k) the threshold $s_0 = 2c$ is nearly always attained, the result of the transformation by the nonlinearity N_{energy} is $kc/2 - s_0$. Additionally, there is an identical contribution to motion energy from the filters (iii) positioned one check behind the moving check. Finally, the net contribution to motion energy for the FM stimulus *R/R*-drift (summed over trajectory positions) is $2(kc/2 - s_0) = kc - 2s_0$.

Thus, for this model, the motion-energy signals generated by corresponding FM and NFM stimuli differ by an amount s_0, which (in the limit that k becomes large) becomes negligible. For both stimuli, the opponent motion energy is zero. This is because every filter (iii) whose space-time orientation is opposite to the direction of stimulus motion receives at most one contribution of size $2c$, and therefore its response is reduced to zero by the threshold component of N_{energy}.

The critical feature of the model was the use of a rectifying nonlinearity, rather than a quadratic one, at the stage of motion-energy calculation. This final nonlinearity N_{energy} may be replaced by a half-wave rectifier without altering the above calculation, provided that the threshold is retained. In this model, the threshold is key to providing direction selectivity, by elim-

inating low-level responses generated by checks moving in the nonpreferred direction. To retain selectivity over a wide range of stimulus contrasts, we postulate that threshold is proportional to contrast.

However, the precise form of the linear filters is not critical to the model. The initial filter (i) must be temporally band-pass or high-pass, so that the presence of changing contrast can be signaled, but its precise form does not matter. Similarly, other spatiotemporal impulse responses for the directionally selective filter (iii) may be used, including, for example, inhibitory sidebands. Properly balanced inhibitory sidebands may reduce the need for the threshold in N_{energy} at stage (iv), since they serve to reduce the response to motion in the nonpreferred direction.

Intuitively, the reason that the rectifier model succeeds in equating responses to FM and NFM while the quadratic model necessarily produces a larger response to FM (for choices of N_{local} with an odd-symmetric component) can be seen by considering the distribution of signals produced by the directionally selective stage (iii). For each stimulus, the array of local nonlinearities produces responses that are equal to 0 or $+2c$. These distributions have equal mean values, but the variance is greater for the FM stimulus than for the NFM stimulus. Since these distributions have the same mean, they produce the same average motion energy as calculated by a rectification. But averaging these signals following a squaring operation produces a larger value in the FM case than in the NFM case, since the distribution in the FM case has a greater variance. This argument shows that rectification at the stage N_{energy} is the only kind of nonlinearity which will (in this model framework) provide for equal responses to FM and NFM.

We therefore propose a two-nonlinearity computational structure for the extraction of motion energy, consisting of a local rectifying subunit and a second stage of thresholded rectification with a broader summing area, and intervening linear stages. A similar structure has been proposed to account for certain aspects of texture perception (Victor & Conte, 1991). As in that model, the second nonlinearity has a threshold proportional to contrast in order to retain specificity over a wide contrast range. This nonlinearity provides for a cooperative interaction among the subunits. Such a cooperative interaction may underlie improved psychophysical performance in motion tasks as the stimulus trajectory is lengthened (McKee & Welch, 1985; McKee & Watamaniuk, 1991).

In the stimuli we have considered, there is only one direction of motion. Stimuli with components or elements moving in multiple directions, such as plaids (Adelson & Movshon, 1982), sheets of random dots, and the phi/reversed phi stimulus of Chubb and Sperling (1989) will generate motion-energy signals from detectors tuned to more than one direction and/or spatial scale. The model as proposed here does not account for the motion percepts elicited by these stimuli, such as transparency, coherence, and capture. An account of these percepts likely rests on understanding how individual motion-energy signals are combined.

A general computational advantage of this model is that it builds in a scaling property—despite the nonlinearities, scaling the contrast of the input scales the contrast of the output by the same factor once threshold is well exceeded. Nonlinear responses generated by this kind of network therefore remain proportional to linear responses over a wide range of suprathreshold contrasts.

Acknowledgments

A portion of this work was presented at the 1990 meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida. This work was supported in part by Grants EY1428 and EY7977.

References

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

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

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

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research* **14**, 519–529.

Braddick, O. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society B* (London) **290**, 137–151.

Cattel, R.B. (1978). *The Scientific Use of Factor Analysis*. New York: Plenum Press.

Chubb, C. & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A* **5**, 1986–2006.

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

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

Emerson, R.C., Citron, M., Vaughn, W.J. & Klein, S. (1987). Nonlinear directionally selective subunits in complex cells of cat striate cortex. *Journal of Neurophysiology* **58**, 33–65.

Gutowitz, H., Zemon, V., Victor, J.D. & Knight, B.W. (1986). Source geometry and dynamics of the visual evoked potential. *Electroencephalography and Clinical Neurophysiology* **64**, 308–327.

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

Julesz, B., Gilbert, E.N. & Victor, J.D. (1978). Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics* **31**, 137–140.

McKee, S. & Watamaniuk, S.N.J. (1991). Detecting a single point moving on a linear trajectory amidst randomly moving points. *Investigative Ophthalmology and Visual Science* (Suppl.) **32**, 892.

McKee, S. & Welch, L. (1985). Sequential recruitment in the discrimination of velocity. *Journal of the Optical Society of America A* **2**, 243–251.

Milkman, N., Schick, G., Rossetto, M., Ratliff, F., Shapley, R. & Victor, J.D. (1980). A two-dimensional computer-controlled visual stimulator. *Behavior Research Methods and Instrumentation* **12**, 283–292.

Nakayama, K. (1985). Biological image motion processing: A review. *Vision Research* **25**, 625–660.

Nakayama, K. & Silverman, G.H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America A* **2**, 267–274.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In *Sensory Communication*, ed. Rosenbluth, W.A., pp. 303–317. New York: John Wiley & Sons.

Simoncelli, E.P. & Adelson, E.H. (1991). Relationship between gradient, spatiotemporal-energy, and regression models for motion perception. *Investigative Ophthalmology and Visual Science* (Suppl.) **32**, 893.

Spekreijse, H., Dagnelie, G., Maier, J. & Regan, D. (1985). Flicker and movement constituents of the pattern reversal response. *Vision Research* **25**, 1297–1304.

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

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

VAN DIJK, B.W., DAGNELIE, G. & SPEKREIJSE, H. (1986). Motion onset-offset evoked potentials from rhesus visual cortex. In *Evoked Potentials III. The Third International Evoked Potentials Symposium*, ed. BARBER, C. & BLUM, T., pp. 205–212. Boston, Massachusetts: Butterworths.

VAN DIJK, B.W. & SPEKREIJSE, H. (1989). Responses to onset and offset of visual motion. *Investigative Ophthalmology and Visual Science* (Suppl.) **30**, 426.

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

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. (1991). Spatial organization of nonlinear interactions in form perception. *Vision Research* **31**, 1457–1488.

VICTOR, J.D., CONTE, M.M. & IADECOLA, C. (1991). Ocular dependence of hemifield visual-evoked potentials: Relevance to bilateral cortical representation of central vision. *Clinical Vision Sciences* **6**, 261–176.

VICTOR, J.D. & MAST, J. (1991). A new statistic for steady-state evoked potentials. *Electroencephalography and Clinical Neurophysiology* **78**, 378–388.