

SPATIAL ORGANIZATION OF NONLINEAR INTERACTIONS IN FORM PERCEPTION*

JONATHAN D. VICTOR and MARY M. CONTE

Department of Neurology and Neuroscience, Cornell University Medical College, 1300 York Avenue and Laboratory of Biophysics, The Rockefeller University, 1230 York Avenue, New York, NY 10021, U.S.A.

(Received 7 February 1990; in revised form 11 January 1991)

Abstract—We examined the perception of structure in a family of visual textures whose second-order correlation structure is flat. These textures were generated by two-dimensional recursion rules, in a manner which extends the construction of Julesz, Gilbert and Victor (1978; *Biological Cybernetics*, 31, 137-140). Textures generated by some recursion rules elicited a visually salient percept of structure, while textures generated by other recursion rules did not. Textures whose statistical structure was visually salient produced evoked responses which differed from the response evoked by completely random textures. The size of this VEP difference correlated well with psychophysical measures.

Since the textures were constructed to have identical global spatial frequency spectra, models for the extraction of visual structure must be essentially nonlinear. Models based on symmetry, information content, or simple spatial extent (but not pattern) of correlation fail to explain the observed results. Models based on the cooperative interaction of pairs of nonlinear subunits provide a reasonable qualitative account of the findings. The critical model features are (i) the presence of multiple nonlinear subunits, and (ii) a second nonlinearity, such as a threshold, at the stage of combination of subunit signals.

Human Visual textures Modeling Nonlinear interactions Visual evoked potentials

INTRODUCTION

The goal of this study is a better understanding of neural computations underlying form analysis. Although many neurons of visual cortex may be classified and described by essentially linear techniques (Hubel & Wiesel, 1968; De Valois, Albrecht & Thorell, 1978) there is much evidence that form analysis is a highly nonlinear process (Movshon, Thompson & Tolhurst, 1978b; Morrone, Burr & Maffei, 1982; Von der Heydt, Peterhans & Baumgartner, 1984; Desimone, Schein, Moran & Ungerleider, 1985).

In a formal sense, nonlinearities signal feature extraction; without nonlinearities, feature extraction and spatial filtering are indistinguishable (Victor, 1988). From a systems-analytic point of view, nonlinearities may serve as clues to the internal structure of a transduction.

Identification of nonlinearities, and separation of linear from nonlinear phenomena, requires a sufficiently rich stimulus set. For this study, we have chosen to use several classes of "isodipole textures" (Julesz et al., 1978). These

are broadband stimuli which contain a variety of local features. One analytical advantage of these stimuli is that isodipole texture pairs (considered as ensembles) have identical spatial frequency spectra. Because of this property, mechanisms which distinguish the stimuli from each other must involve highly nonlinear processes (Victor, 1985; Victor & Zemon, 1985). While (or perhaps because) these textures are unlikely to be encountered in the ordinary visual environment, the mechanisms that they tap must be driven by more familiar stimuli (such as bars and gratings) as well. However, these more familiar stimuli often do not readily allow for a qualitative separation of linear and nonlinear neural mechanisms.

We will explore psychophysical and VEP responses to several classes of isodipole textures, with the goal of analyzing the neural computations underlying early stages of form perception. Previous studies (Victor, 1985; Victor & Zemon, 1985; Victor & Conte, 1989) have allowed us to set lower bounds on the complexity of an adequate model. As we will see, the current studies suggest a model architecture, and simple models within this architecture account for the main features of the data. The critical feature of the architecture is the presence

*A portion of this work was presented at the 1989 Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida.

of two nonlinear stages: a local nonlinear sub-unit, and a long-range cooperative interaction. This stands in contrast to most current models for receptive fields of simple and complex cells (Movshon et al., 1978a, b; Spitzer & Hochstein, 1985a, b; Camarda, Peterhans & Bishop, 1985; Emerson, Citron, Vaughn & Klein, 1987; Jones & Palmer, 1987; Szulborski & Palmer, 1990), which feature at most a single nonlinear stage.

METHODS

Visual stimuli

All of the visual stimuli used in this study are constructed out of pairs of isodipole textures. Two textures are said to be "isodipole" (Julesz et al., 1978) if their first-order and second-order statistics agree. The first-order statistic is simply mean luminance; the second-order statistic is the autocorrelation function. Despite the apparently restrictive nature of this defining condition, a number of algorithms exist for generation of isodipole textures. The first stage of the construction uses Gilbert's (1980) algorithm to generate many classes of highly-ordered isodipole textures. The second stage of the construction is an extension of that of Diaconis and Freedman (1981), which allows for graded levels of structure within each texture class.

The reader is reminded that the motivation for the use of statistically complex visual stimuli is to separate nonlinear processes from spatial filtering and to make inferences as to the nature of the nonlinearity. As pointed out in the Discussion, our ability to resolve the presence of two nonlinear stages would be lost if we restricted our stimuli to bars, gratings, or other traditional stimuli.

Classes of isodipole textures. Gilbert's (1980) algorithm is essentially a generalization of the algorithm of Julesz et al. (1978) used to generate the basic "even" texture used in our previous work. The algorithm assigns a state of +1 or -1 to each check of a lattice in a recursive fashion. The state assigned to the check in row *i* and column *j* will be denoted by $a_{i,j}$. For the original algorithm of Julesz et al., the recursion rule is:

$$a_{i,j} = a_{i-1,j} \cdot a_{i,j-1} \cdot a_{i-1,j-1}. \tag{1}$$

This rule requires that there are an even number of checks whose state $a_{i,j} = 1$ within any 2×2 -region of the texture. The rule (1) may be written more conveniently as:

$$\prod_{(I_s, J_s) \text{ in } G} a_{i+I_s, j+J_s} = 1, \tag{2}$$

where the set *G* contains the coordinates of four points in a standard 2×2 windowpane: $G = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

Following Gilbert (1980), we generalize this rule by replacing the 2×2 -region *G* by regions of other shapes. We will call a region *G* which generates a texture according to equation (2) a *glider*. A texture created with the generalized recursion rule (2) will have an even number of checks whose state $a_{i,j} = 1$ within any region whose shape is that of the glider *G*. The gliders *G* we have used are illustrated in Table 1, and the textures they generate are shown in Fig. 1. A texture created with a glider *G* will be denoted T_G .

There are two technical points worth mentioning. (i) The recursion rule (2) only suffices to generate textures once the states of some of the checks have been initialized in a random, uncorrelated fashion. It might at first seem that the statistics of the texture that is generated

Table 1. Construction of selected isodipole textures. The coordinates of the glider are used in the recursion rule (2). For the glider sets and diagrams, the first coordinate (*x*) indicates the row (from bottom to top), and the second coordinate (*y*) indicates the column (from left to right)

Texture type	Glider coordinates (row, column)	Glider shape
Standard	{(0, 0), (0, 1), (1, 0), (1, 1)}	<pre> • • • x x • x x • </pre>
Cross	{(0, 1), (1, 0), (1, 2), (2, 1)}	<pre> • x • x • x • x • </pre>
Zigzag	{(0, 0), (0, 1), (1, 1), (1, 2)}	<pre> • • • • x x x x • </pre>
Oblong	{(0, 0), (0, 2), (1, 0), (1, 2)}	<pre> • • • x • x x • x </pre>
Triangle	{(0, 1), (1, 0), (1, 1)}	<pre> • • • x x • • x • </pre>
Tee	{(0, 0), (0, 1), (0, 2), (1, 1)}	<pre> • • • • x • x x x </pre>
Wye	{(0, 2), (1, 0), (1, 1), (2, 1)}	<pre> • x • x x • • • x </pre>
Foot	{(0, 0), (0, 2), (1, 1), (1, 2)}	<pre> • • • • x x x • x </pre>
El	{(0, 2), (1, 0), (1, 1), (1, 2)}	<pre> • • • x x x • • x </pre>

might depend on the choice of the initial set. It is also *a priori* possible that the statistics of the texture might depend on the manner in which the recursion rule is applied (e.g. row first, or column first). However, Gilbert (1980) has shown that neither of these potentially complicating issues can arise; the statistics of a texture generated by equation (2) depend only on the glider G , and not on the initial set or the direction of propagation. In practice, we have found it convenient to initialize one or more rows and one or more columns, and to generate the texture row by row. The number of rows and columns which need to be initialized depends on the texture class to be generated. (ii) For textures generated by n -element gliders, one anticipates an n th-order correlation between

checks, corresponding to the rule (2). However, for some regions G , correlations of lower order may also be nonzero (for example, a glider consisting of several consecutive checks in one row). Gilbert (1980) provided a means to determine which gliders G lead to *no* correlations at order less than n ; we have only used gliders which conform to his criteria. Thus, the textures described in Table 1 and illustrated in Fig. 1 all have first- and second-order statistics which are identical to that of a random coloring, and the textures generated with four-element gliders also have third-order statistics which are identical to that of a random coloring.

Decorrelation. In the second stage of the construction, we introduce spatial decorrelation into the textures T_G . As in the previous study

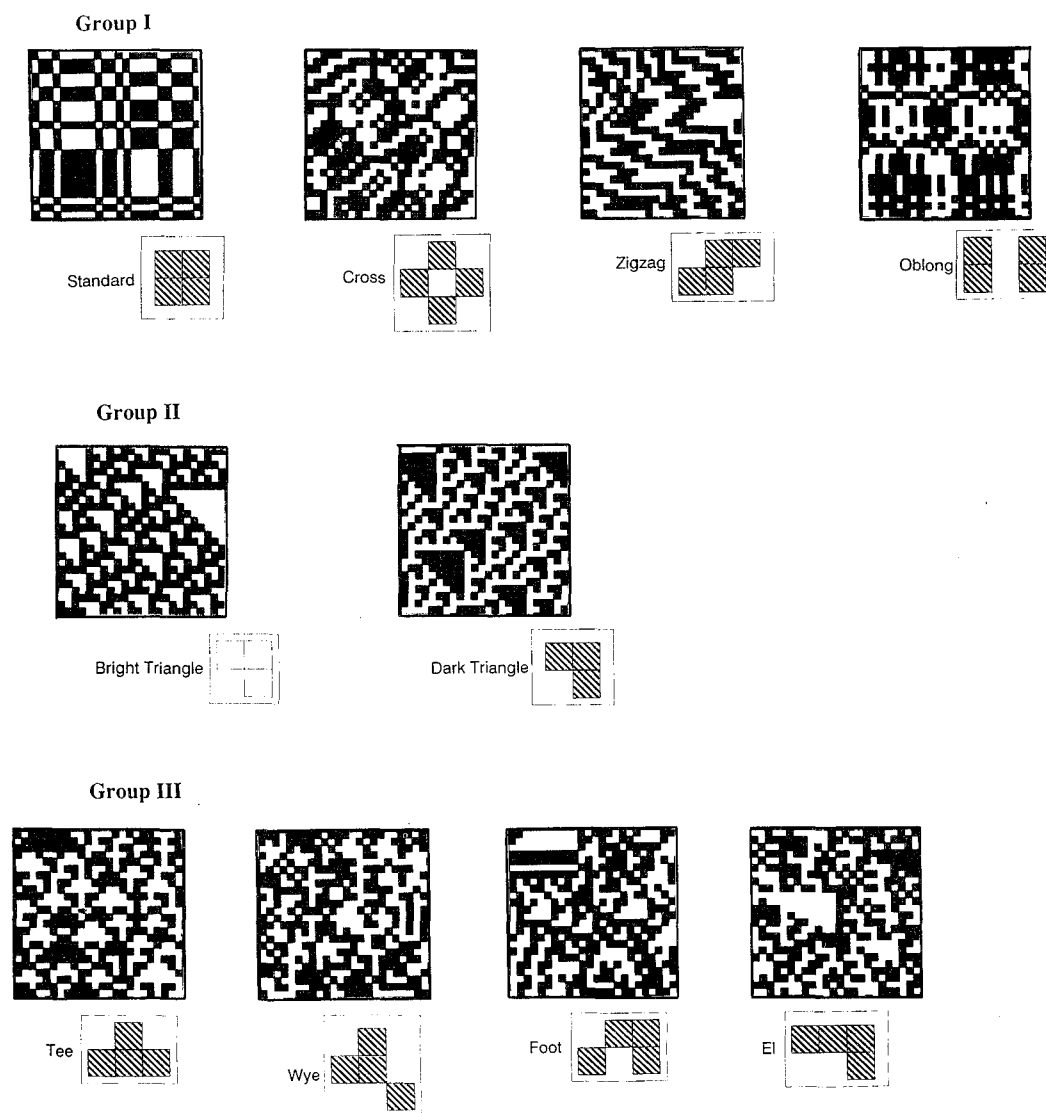


Fig. 1. Examples of visual stimuli used in this study. These textures are fully-correlated textures $T_G(1, 1)$ based on the gliders G of Table 1.

(Victor & Conte, 1989), we consider two kinds of disorder: *propagated* disorder, parametrized by ϵ_{prop} ; and *sporadic* disorder, parametrized by ϵ_{spor} .

To introduce disorder which is propagated throughout the texture, the deterministic recursion rule (2) is replaced by a stochastic one:

$$\text{prob} \left\{ \prod_{(i_s, j_s) \text{ in } G} a_{i+I_s, j+J_s} = 1 \right\} = 1 - \epsilon_{\text{prop}}. \quad (3)$$

Errors may also be introduced in a purely local ("sporadic") fashion. Let us denote the appearance of a check in row i and column j by $A_{i,j}$, where $A_{i,j} = +1$ denotes a bright check, and $A_{i,j} = -1$ denotes a dark check. Thus, the contrast of the check in row i and column j is equal to $A_{i,j}$ multiplied by an overall contrast value. The simplest way to translate states $a_{i,j}$ into appearances $A_{i,j}$ is that the appearance of a check is faithfully determined by its state, i.e. $A_{i,j} = a_{i,j}$. Sporadic errors may be introduced by nonfaithful rendition of the states $a_{i,j}$ into appearances $A_{i,j}$, with the rate of this nonfaithful rendition determined by ϵ_{spor} :

$$\text{prob} \{ A_{i,j} = a_{i,j} \} = 1 - \epsilon_{\text{spor}}. \quad (4)$$

Sporadic decorrelation is one of the processes discussed by Diaconis and Freedman (1981).

In describing the correlation structure of these textures, we have found it useful to introduce the correlation parameters c_{prop} and c_{spor} (Victor & Conte, 1989). These quantities are related to the error rates ϵ_{prop} and ϵ_{spor} by:

$$c_{\text{prop}} = 1 - 2\epsilon_{\text{prop}} \quad (5)$$

and

$$c_{\text{spor}} = (1 - 2\epsilon_{\text{spor}})^4. \quad (6)$$

A texture constructed from error parameters ϵ_{prop} and ϵ_{spor} and glider G will be denoted by $T_G(c_{\text{prop}}, c_{\text{spor}})$, where c_{prop} and c_{spor} are given by equations (5) and (6). In this notation, the textures illustrated in Fig. 1 are textures $T_G(1, 1)$, for the gliders G listed in Table 1. The texture $T_{\text{standard}}(1, 1)$ is the even texture of Julesz et al. (1978). Several sporadically-decorrelated textures $T_{\text{zigzag}}(1, \alpha)$ are illustrated in Fig. 2. Note that the textures $T_G(c_{\text{prop}}, 0)$ and $T_G(0, c_{\text{spor}})$ are statistically identical to the completely random texture $T(0, 0)$ for any glider G , and for any value of c_{prop} and c_{spor} .

SPORADIC DECORRELATION

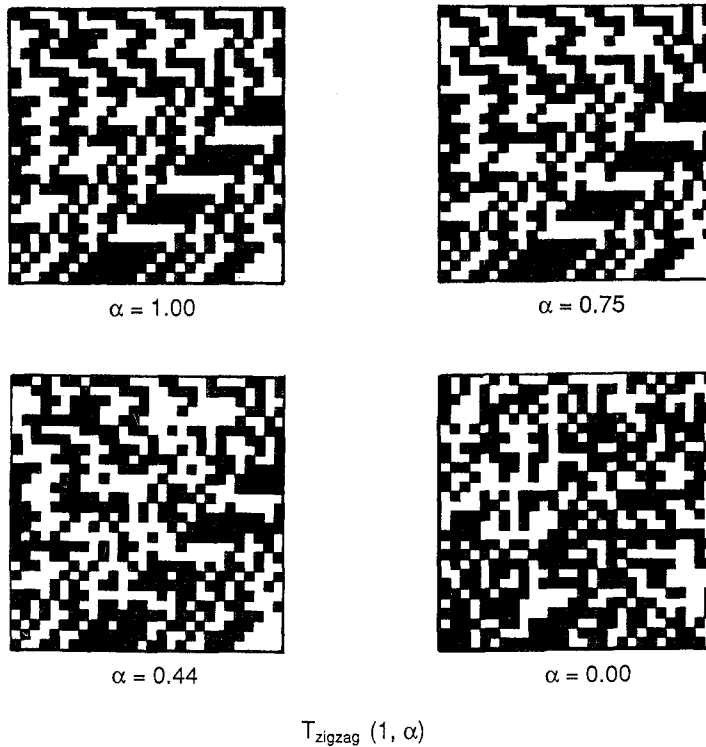


Fig. 2. Examples of sporadically-decorrelated textures $T_{\text{zigzag}}(1, \alpha)$ based on the zigzag glider of Table 1. The texture $T_{\text{zigzag}}(1, 0)$ is fully decorrelated (i.e. completely random).

Correlation statistics. We next state the correlation statistics of the textures $T_G(c_{prop}, c_{spor})$. To do this, we need to consider superpositions of the glider G and translations of itself. Following Gilbert (1980), we introduce a generating function $\hat{S}(x, y)$ associated with a set of checks $S = \{(I_s, J_s)\}$:

$$\hat{S}(x, y) = \sum_{(I_s, J_s) \text{ in } G} x^{I_s} y^{J_s}. \quad (7)$$

The generating function has several useful properties. Translation of a set S by one unit along the x -axis corresponds to multiplication of \hat{S} by x , and translation along the y -axis corresponds to multiplication of \hat{S} by y . Superposition of sets S_1 and S_2 corresponds to addition of their generating functions \hat{S}_1 and \hat{S}_2 .

It will be convenient to consider generating functions as algebraic expressions in x and y which can be multiplied and added with all the usual rules for multiplication of polynomials, with the coefficients interpreted modulo 2 (i.e. in the finite field of two elements $\{0, 1\}$). Multiplication of two generating functions \hat{S}_1 and \hat{S}_2 now has the following interpretation: the product $\hat{S} = \hat{S}_1 \hat{S}_2$ is the generating function of a set S whose members (I, J) can be written in an *odd* number of ways as a vector sum of a member (I_1, J_1) of S_1 and a member (I_2, J_2) of S_2 . Multiplication of generating functions thus corresponds to convolution (mod 2) of sets.

We say that the set S is "divisible" by the glider G if there exists a finite quotient set Q whose generating function \hat{Q} satisfies:

$$\hat{S}(x, y) = \hat{G}(x, y) \cdot \hat{Q}(x, y) \pmod{2}. \quad (8)$$

If S is divisible by G and Q is the smallest set satisfying equation (8), then we will write $Q = S/G$. The number of elements in Q will be denoted by $n(Q) = n(S/G)$.

Here is an example. The generating function $\hat{S} = (1 + x^I)(1 + y^J) = 1 + x^I + y^J + x^I y^J$ corresponds to the four corners of a rectangle of dimension $(I + 1) \times (J + 1)$. The generating function $\hat{G} = (1 + x)(1 + y) = 1 + x + y + xy$ corresponds to the glider for the standard even texture. In this case, \hat{S} is divisible by \hat{G} and the quotient polynomial is given by:

$$\hat{Q} = (1 + x + \dots + x^{I-1}) \times (1 + y + \dots + y^{J-1}). \quad (9)$$

\hat{Q} is thus seen to contain IJ terms $x^i y^j$, one corresponding to each of the checks in an $(I - 1) \times (J - 1)$ rectangle (edges and interior). So in this case, $n(S/G) = IJ$.

The notion of divisibility permits a concise statement of the correlation properties of the textures $T_G(c_{prop}, c_{spor})$. For checks positioned at a set of coordinates S which is *not* divisible by G , all correlations are equal to zero. For checks positioned at a set of coordinates S which *is* divisible by G , correlations are given by:

$$\left\langle \prod_r A_{i+I_r, j+J_r} \right\rangle = (c_{spor})^{n(S)/4} \cdot (c_{prop})^{n(S/G)}. \quad (10)$$

These results may be verified by the techniques of Appendix 1 of Victor and Conte (1989), and Theorem 2 of Gilbert (1980) for fully-ordered textures $T_G(1, 1)$.

The above analysis guarantees that for textures $T_G(c_{prop}, c_{spor})$, all correlations among $n(G) - 1$ or fewer checks are zero. {In the language of Victor (1988), these textures are iso- $[R^*, n(G) - 1]$ for all receptive field profiles R^* }. It also provides a means for construction of distinct textures $T_G(1, c_{spor})$ and $T_G(c_{prop}, 1)$ for which correlations across checks within the glider set G are identical. For one of these textures $[T_G(c_{prop}, 1)]$, long-range correlations diminish to zero; for the other $[T_G(1, c_{spor})]$, long-range correlations do not diminish to zero.

This analysis also shows that there is a qualitative difference between textures constructed on gliders of size 4 (e.g. the standard texture) and gliders of size 3 (e.g. the triangle textures). For textures whose generating glider has an even number of elements, contrast-inversion merely provides a new example of the same texture class. But for textures whose generating glider has an odd number of elements, contrast-inversion leads to a statistically distinct texture. To see this, observe that contrast-inversion of a texture is equivalent to choosing $\epsilon_{spor} = 1$: the states of all checks are inverted. For textures with $n(G)$ even, nonzero correlations exist on a set S only if $n(S)$ is even as well (Gilbert, 1980). Hence, $n(S)$ must be even in equation (10), and $\epsilon_{spor} = 1$ leads to the same correlation structure as does $\epsilon_{spor} = 0$. However, for textures with $n(G)$ odd, some sets S with an odd number of elements (for example, $S = G$) do lead to nonzero correlations. For these sets, contrast-inversion inverts the sign of the correlation.

There is a technical comment concerning the definition of the sporadic correlation c_{spor} (equation 6). We will be comparing psychophysical and electrophysiological responses to textures with $n(G) = 4$ and $n(G) = 3$, and we will use $\alpha = c_{spor}$ as the independent variable for

many of these comparisons. For textures with $n(G) = 4$ or $n(G) = 3$, c_{spor} indicates the strength of correlations among elements of sets S of size 4. However, for the textures with $n(G) = 3$, there are additional correlations of strength $(1 - 2c_{\text{spor}})^3 = \alpha^{3/4}$ among elements of sets S of size 3 [i.e. $S = G$ in equation (10)]. Such third-order correlations are not present in textures with $n(G) = 4$.

General experimental procedures. The stimuli described above were produced on a Tektronix 608 monitor. Control signals for the raster display [horizontal (X), vertical (Y), and intensity (Z)] were generated by specialized electronics (Milkman, Schick, Rossetto, Ratliff, Shapley & Victor, 1980) interfaced to a DEC 11/73 computer. The display subtended 8.8 deg and had a luminance of 150 cd/m². Stimuli had a contrast $[(I_{\text{max}} - I_{\text{min}})/(I_{\text{max}} + I_{\text{min}})]$ of 0.4, and were viewed binocularly at a distance of 57 cm.

The subject pool for VEP and psychophysical studies consisted of six normal volunteers, two male and four female; four of these subjects were naive to the purpose of the experiments. Subject CI is a protanope as determined by Ishihara and Farnsworth–Munsell 100-Hue (Farnsworth, 1957) tests. Subjects ranged in age from 22 to 40 yr, and were corrected to normal visual acuity if necessary.

Psychophysical methods. After a subject-paced warning sound, a stimulus would appear from an equiluminous uniform background for a period of 50 msec. This stimulus had a 50% probability of being a random stimulus $T(0, 0)$, and a 50% probability of being a structured stimulus $T_G(1, c_{\text{spor}})$. Then, an uncorrelated random mask $T(0, 0)$ of equal contrast and luminance would appear for 500 msec. The observer's forced-choice task was to distinguish the structured stimulus from a random one.

For each stimulus type listed in Table 1, the independent variable was the level of sporadic decorrelation c_{spor} . Trials were presented in blocks of 20 of a particular stimulus type and value of c_{spor} . Before data collection, observers were given practice (with feedback) with free-view and timed presentations of examples of the target textures and random textures. Practice was allowed until performance stabilized; no feedback was given during data collection. A set of 20 blocks (two repeats of each of 10 values of c_{spor} for a particular glider G) were presented in random order once performance had stabilized. In one testing session, four or five sets of 20-stimulus blocks were presented. For each

subject, each glider was tested on more than one day, so that at least four blocks (80 trials) were collected for most stimulus types and values of c_{spor} . However, in order to minimize collection of redundant data, only two or three blocks were collected for stimuli for which performance was perfect at a lower level of correlation.

Measurement of visual evoked potentials. Visual evoked potentials (VEPs) were elicited by periodic abrupt interchange between examples of two texture classes. The stimulus cycle lasted 474 msec. In the first half of each stimulus cycle (237 msec), an example of a particular structured texture $T_G(c_{\text{prop}}, c_{\text{spor}})$ was presented. In the second half of each stimulus cycle, an example of a random texture $T(0, 0)$ was presented. In each successive stimulus cycle, new examples of the structured texture $T_G(c_{\text{prop}}, c_{\text{spor}})$ and the random texture $T(0, 0)$ were presented. The VEP was averaged over this stimulus cycle, so that it represented an averaged response to many examples of each texture class. At each stimulus transition, each check of the texture had a 50% probability of reversing contrast. There was no correlation between the positions of the checks that reversed in successive stimulus transitions. Each raw waveform represents an average of responses recorded in two or three one-minute runs. These runs were presented in randomized blocks.

Scalp signals were recorded differentially at C_z-O_z , P_z-O_z , C_z-O_1 , C_z-O_2 , amplified 10,000-fold, and bandpass-filtered (0.03–100 Hz) prior to digitization. Artifact rejection was performed by the computer prior to averaging. VEPs were saved by the 11/73 computer for offline analysis, which included Fourier analysis, modeling, and factor analysis.

RESULTS

Psychophysics

Inspection of the isodipole textures (Fig. 1) revealed that for some textures, the underlying structure was readily apparent (e.g. the standard texture). For other textures, the underlying structure was not readily discernable (e.g. the wye texture). We first quantify the psychophysical salience in the structured textures by a forced-choice psychophysical paradigm, and then introduce a parametrization of the psychophysical curves to allow comparison of performance within and across subjects.

Psychometric functions. Stimuli were presented for 50 msec, followed by a random mask. The

subject was required to discriminate between examples of a sporadically-decorrelated structured texture and examples of random textures. For each texture, the task was made progressively more difficult by reducing the degree of local correlation $c_{spor} = \alpha$ from 1 to 0.

Results for two subjects are shown in Fig. 3. For some textures, discriminability was essentially perfect for $\alpha = 1$ (no decorrelation), and remained good as α decreased. These textures included four of the four-element glider textures, as well as the light and dark polarities of the triangle texture. Below, we will refer to the four-check glider textures for which structure is readily apparent as "Group I," and the three-check glider textures for which structure

is readily apparent as "Group II." That is, Group I consists of the standard, cross, zigzag, and oblong textures, and Group II consists of the bright-triangle and dark-triangle textures. For these textures, rapid discriminability remained above chance even when local correlation is small ($\alpha = 0.1$).

For the other textures, discriminability was imperfect even for $\alpha = 1$. Discriminability deteriorated further as decorrelation was introduced ($\alpha < 1$), usually becoming close to chance for $\alpha \leq 0.3$. These textures (tee, wye, foot, and el textures) will be referred to as "Group III" textures.

Extraction of psychophysical parameters. In order to compare performance within and

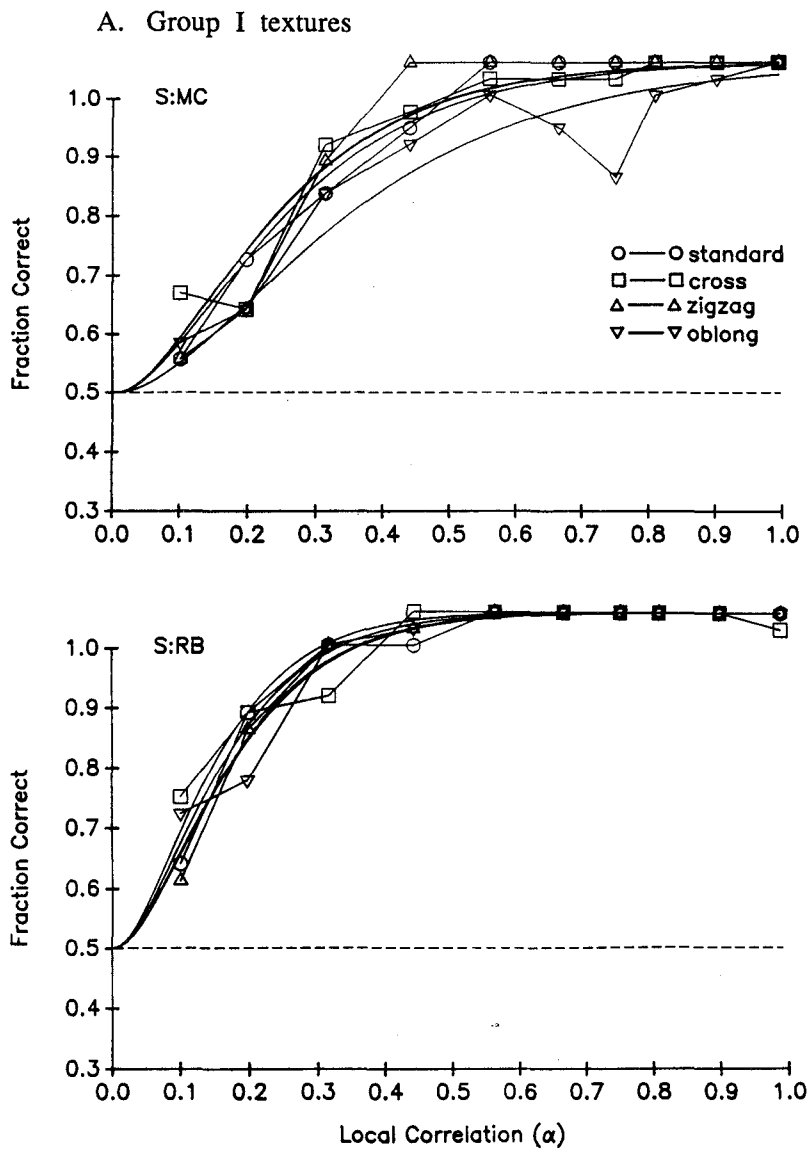


Fig. 3(A). Caption on p. 1465.

across subjects, we used a simple psychophysical model to parametrize these curves. This psychophysical model postulates that there are N independent detectors available to detect the presence of structure. At the beginning of the stimulus presentation, all detectors are in an "initial", or undecided, state. During stimulus presentation, detectors independently convert into a "set" state. At the end of the stimulus presentation, if the number of detectors in a "set" state exceeds a criterion N_c , then the subject performs correctly. If not, then the subject guesses, with a 50% probability of being correct. We postulate that the rate at which detectors convert is the product of two factors: the local correlation α and a rate constant r .

The three parameters of the model (the rate constant r , the critical number of detectors N_c , and the number of detectors N) are allowed to depend on the glider type, G .

This psychophysical model is analyzed in detail in Victor and Conte (1990). An important result is that the model behavior depends primarily on the critical number of detectors, N_c , and their population rate of conversion, $R = Nr$. The absolute number of detectors N has little influence on model behavior. Parameters N_c and R of the model were determined for the psychophysical performance of each of six subjects for each of the 10 texture types.

Examples of fits of the model for $N_c = 2$ to experimental data are shown as the smooth

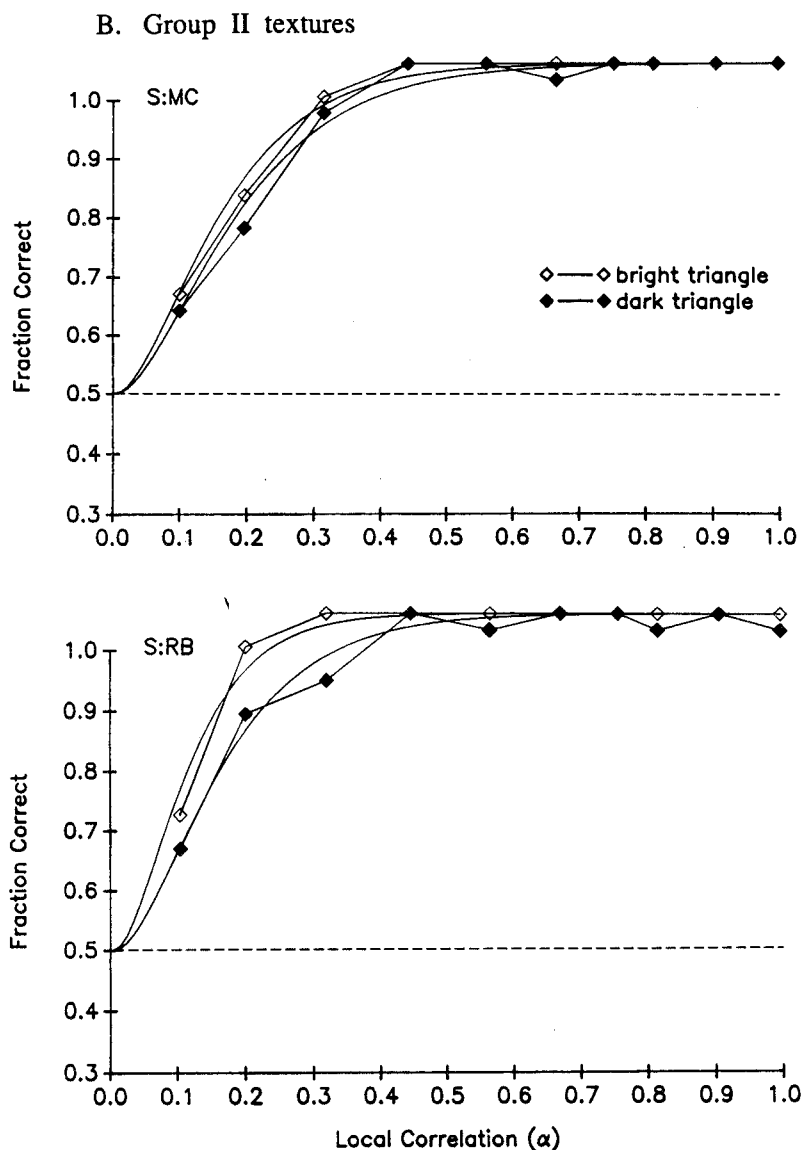


Fig. 3(B). Caption on facing page.

curves in Fig. 3. As seen in the figure, quite reasonable fits were obtained for Group I and Group II textures. The fits for data from Group III textures were not as good, but this is not surprising given the greater internal variability of these data (Fig. 3C). Over all groups of textures and subjects, model and measured performance typically differed by an r.m.s. of 0.03. This is a good fit, considering that the uncertainty in the estimate of psychophysical performance as assessed in 80 trials is on the order of 0.1.

Overall, the choice $N_c = 2$ provided the best fits, as judged by the r.m.s. deviation between predicted and measured performance. For some subjects under some conditions, higher values or lower values of N_c provided a smaller value of the r.m.s. deviation. However, the r.m.s. improvement provided by $N_c = 1$ or $N_c \geq 4$ was typically only 0.01. We do not mean to imply by this discussion that the model itself is correct or that $N_c = 2$ has a particular physiologic significance. Rather, we use the model to provide a single parameter $R = R(G)$ (determined by a

C. Group III textures

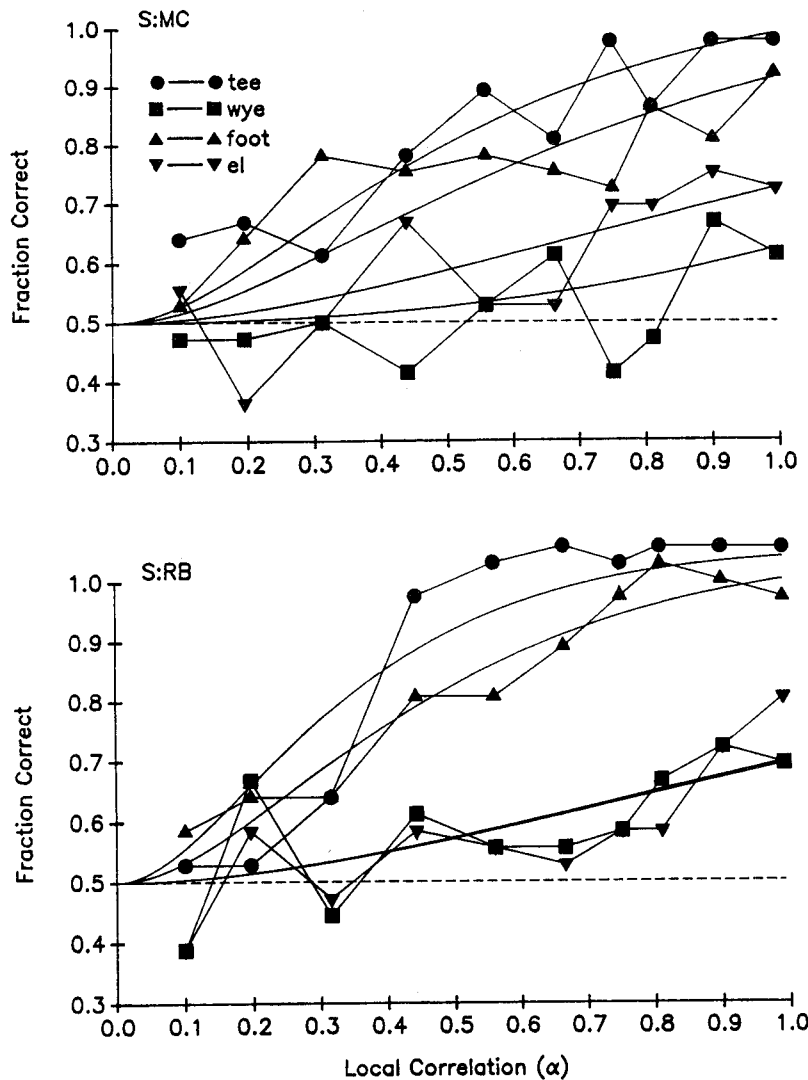


Fig. 3(C).

Fig. 3. Frequency of correct identification of structured isodipole textures in a forced-choice psychophysical task, as a function of local correlation α . The smooth curves represent the model fit (Victor & Conte, 1990) with $N_c = 2$. Part A: Group I textures. Part B: Group II textures. Part C: Group III textures. Subjects: MC, RB.

least-squares method with $N_c = 2$) to compare performance within and across subjects.

Values of the rate constant $R(G)$ computed in this fashion are shown in Table 2. Although there are variations of more than a factor of two in the size of the rate constant R across subjects, there are a number of important consistencies. Rate constants for Group I and Group II textures were greater than those for Group III textures, usually by a factor of four or more. Generally, rate constants for Group II textures exceeded those for Group I textures, but there was overlap for three subjects and the average rate constant for Group II textures was only approx. 35% higher than for Group I textures.

Within groups, rank orders of rate constants were similar across subjects. In Group I, highest rate constants were found for the "standard" or "zigzag" textures in five of six subjects, and lowest rate constants for the "cross" or "oblong" textures in all six subjects. In Group III, although performance was poor, all subjects showed the same rank-order for rate constants: tee > foot > el > wye. In two of the subjects, the best-fitting value of the rate constant R was zero for the "wye" texture. This corresponds to psychophysical performance at chance levels even for the fully-correlated texture ($\alpha = 1$). However, there was no consistent difference in rate constants in Group II textures: rate constants were faster for "bright triangles"

in two subjects, faster for "dark triangles" in three subjects, and nearly identical in one subject. Across subjects, the geometric mean of the rate constants for Group II textures differed by only 7%.

Evoked potentials

We now turn to analysis of VEPs elicited by interchange of the various structured textures with random textures. First we illustrate general features of this VEP and its dependence on stimulus type and local correlation α . Then, we focus separately on its temporal and spatial aspects.

General description. VEPs elicited by interchange between structured isodipole textures and random textures elicited similar waveforms at all derivations examined

$$(C_z-O_z, P_z-O_z, C_z-O_1, C_z-O_2).$$

Generally, responses at P_z-O_z were approximately half the amplitude of the response at C_z-O_z , and there were no consistent differences between the midline derivation C_z-O_z and lateral derivations C_z-O_1 and C_z-O_2 . Responses for one subject at the midline derivation C_z-O_z are shown in Fig. 4. Responses contained a prominent occiput-positive (downward) deflection approx. 100 msec after each transition. For some texture stimuli, such as the standard texture, there is a prolonged occiput-negative (upward) deflection following the transition

Table 2. Rate constants $R(G)$ (sec^{-1}) which provide best fits to observed psychophysical data with $N_c = 2$. The parameter $R(G)$ represents the population rate at which hypothetical "detectors" respond to the isodipole texture G , and is equal to the rate r at which an individual detector responds, times the number of detectors N . For further details, see text. In calculating geometric means, values of zero were excluded

Texture type	Subject						Geometric mean
	JV	MC	SU	GG	CI	RB	
Group I							
Standard	409	142	212	204	161	228	212
Cross	213	148	118	123	110	232	151
Zigzag	234	153	205	244	139	216	201
Oblong	240	108	169	109	96	215	147
Group mean	264	137	172	161	124	223	175
Group II							
Bright triangle	251	223	246	243	193	326	244
Dark triangle	298	195	353	158	192	227	228
Group mean	273	206	295	196	193	272	236
Group III							
Tee	55	71	95	84	86	109	81
Wye	10	13	0	19	0	24	16
Foot	52	53	62	52	44	79	56
El	38	27	61	26	33	25	33
Group mean	32	34	71	38	50	48	39

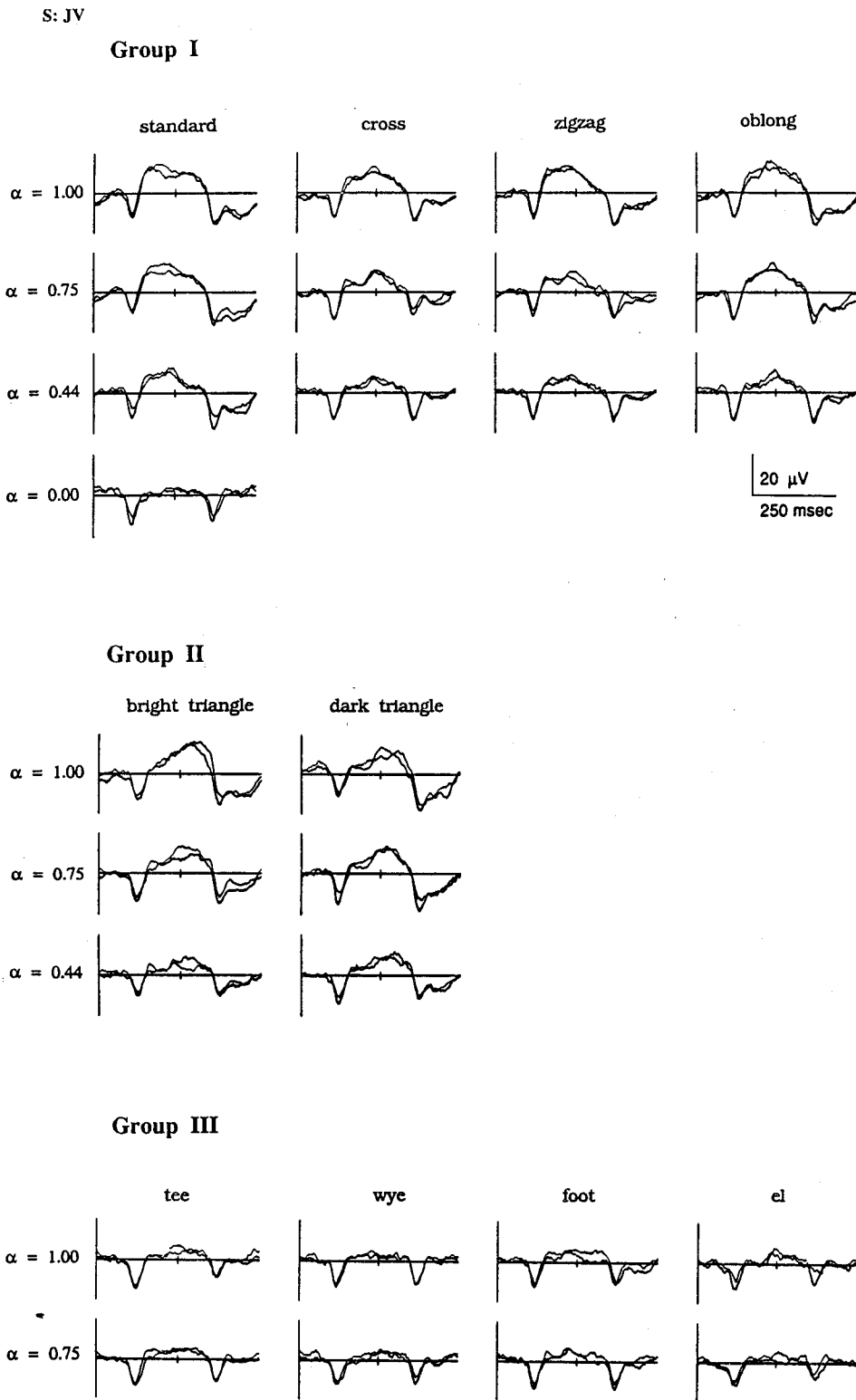


Fig. 4. VEPs elicited by interchange between sporadically-decorrelated isodipole textures and the random texture, as a function of local correlation α . Potentials which are positive at O_2 relative to C_2 are plotted below the isopotential line. Responses to interchanges of examples of the random texture are illustrated as the " $\alpha = 0$ " response to the standard texture in Group I. In each record, the two traces represent the average of separate 1-min runs. Subject: JV.

to the structured texture, which is absent following the transition to the random texture. This feature is seen in responses to textures of Groups I and II but not for responses to the textures of Group III.

We analyze VEP responses to isodipole interchange by decomposing them into two components: a "symmetric" component and an "antisymmetric" component (Victor, 1985; Victor & Zemon, 1985). The symmetric component is the sum of the even harmonics of the VEP waveform. Since the symmetric component is effectively the average of the response to a structured texture and the random texture, it contains responses of neural elements sensitive to local luminance and contrast changes. It is thus analogous to the traditional P-100 (Regan, 1989). As seen in Fig. 5A, the symmetric component is generally similar in voltage and waveform for all of the textures studied. The symmetric component may contain pattern-dependent responses as well; these will be separated below.

The antisymmetric component is the sum of the odd harmonics of the VEP waveform. It is effectively the difference between the response to the onset of the structured texture and the response to onset of the random texture. Since structured and random textures are equated for luminance, contrast, and global spatial frequency content, the antisymmetric component contains responses of neural elements sensitive to more complex aspects of local pattern (Victor, 1985). As seen in Fig. 5B, antisymmetric components are large for the textures of Groups I and II, but small to undetectable for the textures of Group III. Within each group, there is variation in the size of the antisymmetric component driven by different structured textures. To a first approximation, these correspond to relative psychophysical distinguishability. For the subject illustrated in Fig. 5B, the largest Group I response was to the standard texture; this texture also yielded the largest rate constant (Table 2). Similarly, just-detectable antisymmetric components were present for all Group III textures except the "wye" texture; the "wye" texture also was the most difficult one to detect (smallest value of R) psychophysically.

In addition to overall amplitude differences between antisymmetric components elicited by the textures, there is a suggestion that the time-course of the response is different as well. For example, the antisymmetric component

driven by the standard texture is not simply a sinusoid; it has a "nose" on its ascending limb. This nose is not as prominent in the antisymmetric components driven by the other Group I textures, or by the triangle textures.

To look for additional evidence that the pattern-dependent VEP response varies in waveform as well as amplitude, we subjected the symmetric component to a further analysis. If the symmetric component depended merely on local luminance and contrast changes, then an identical symmetric component would be elicited by interchange of any of the isodipole textures, including interchange among examples of the random texture. Thus, any pattern-dependence within the symmetric component elicited by structure texture/random texture interchange may be isolated by subtracting the symmetric component elicited by random texture/random texture interchange.

The results of this calculation, which we will call the "pattern-dependent symmetric subcomponent," are shown in Fig. 5C. For Groups I and III textures, there is little detectable pattern-dependence revealed by this technique. However, data from both Group II textures show a clear dependence on pattern, as manifest by a symmetric subcomponent of approx. $3 \mu\text{V}$. The presence of a pattern-dependent symmetric subcomponent for Group II textures, but not for Group I textures, is evidence that the pattern-dependence of VEP responses is not simply one of amplitude changes.

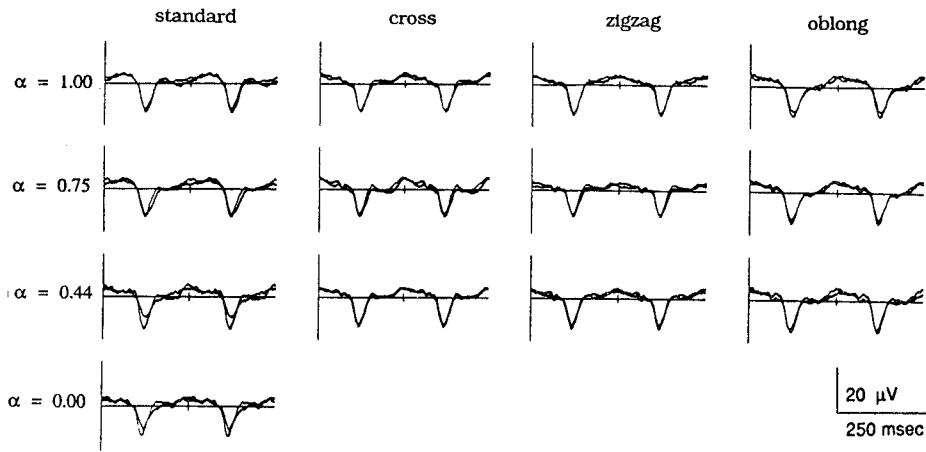
Note that there would be nothing to gain by applying this subtraction technique to the antisymmetric component, since the antisymmetric response to interchange of random textures is essentially zero (Fig. 5B).

VEP responses: overall pattern-dependent amplitudes. The above analysis reveals that the pattern-dependent VEP responses to a range of textures differ not just in overall amplitude, but also in dynamics. However, before we pursue a more detailed analysis of the dynamics, we first summarize the overall size of the pattern-dependent VEP response by the amplitude of the first harmonic, the main contributor to the antisymmetric component.

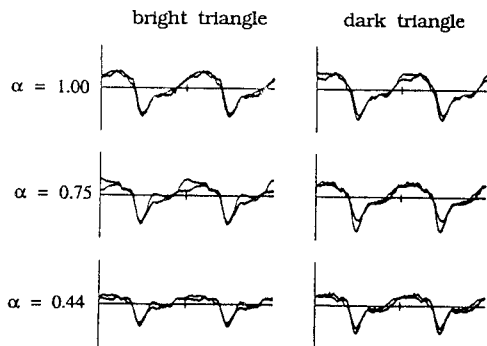
Previously (Victor & Conte, 1989), we showed that the first harmonic amplitude was approximately linearly-related to the local correlation α and that phase was relatively independent of α . This relationship held for responses to the more general isodipole textures examined here, although the proportionality between local

A. Symmetric Component

S:JV Group I



Group II



Group III

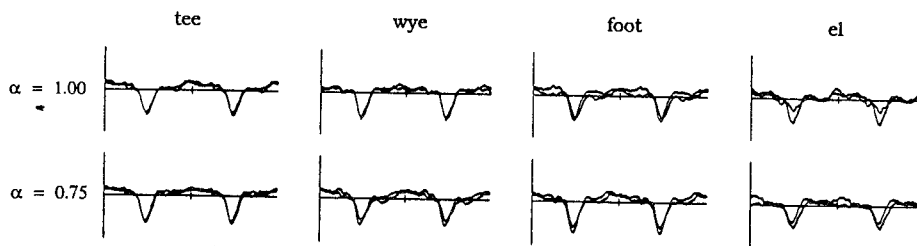
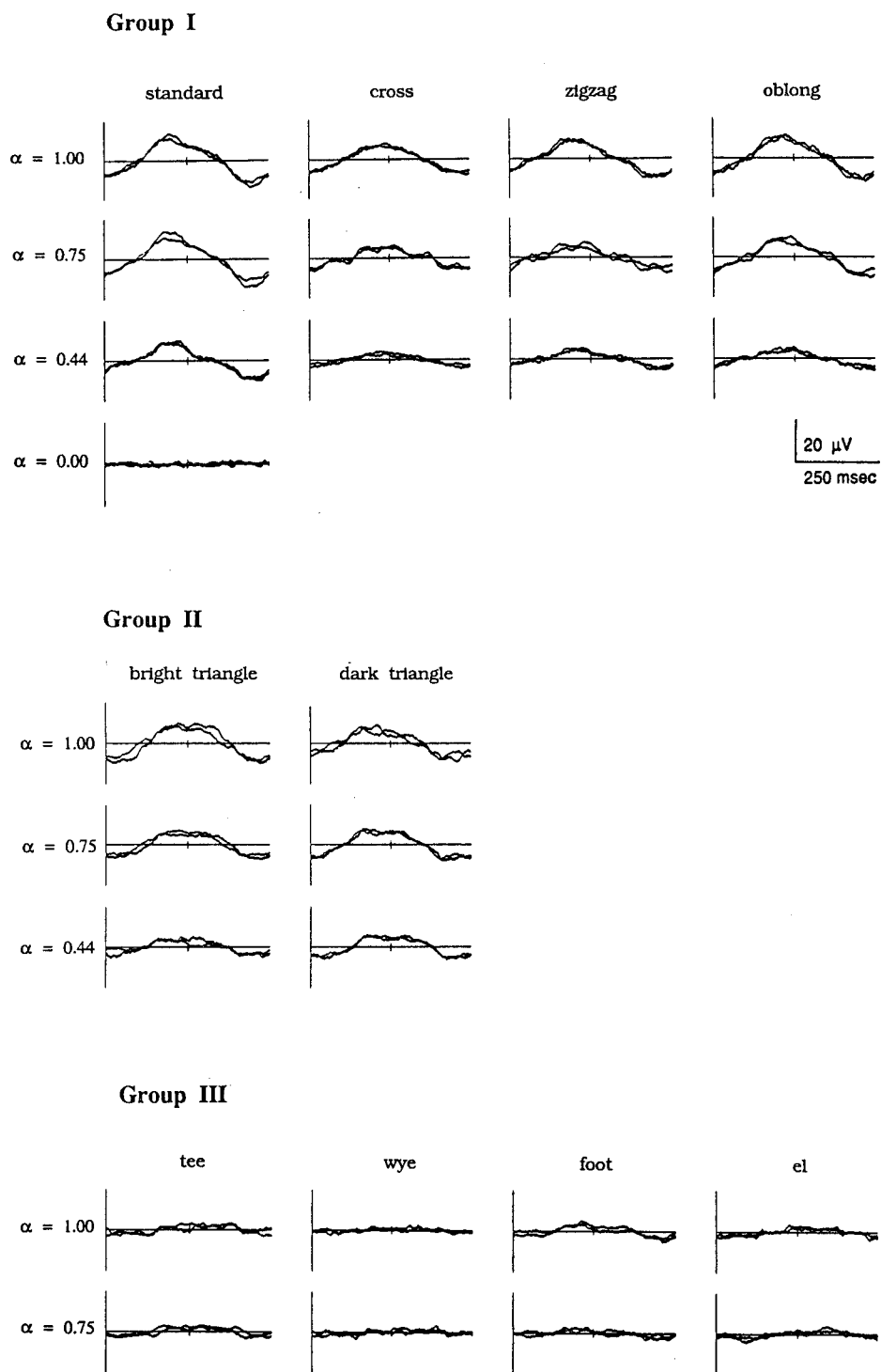


Fig. 5(A). Caption on p. 1471.

B. Antisymmetric Component

Fig. 5(B). *Caption on opposite page.*

C. Pattern-dependent Symmetric Subcomponent

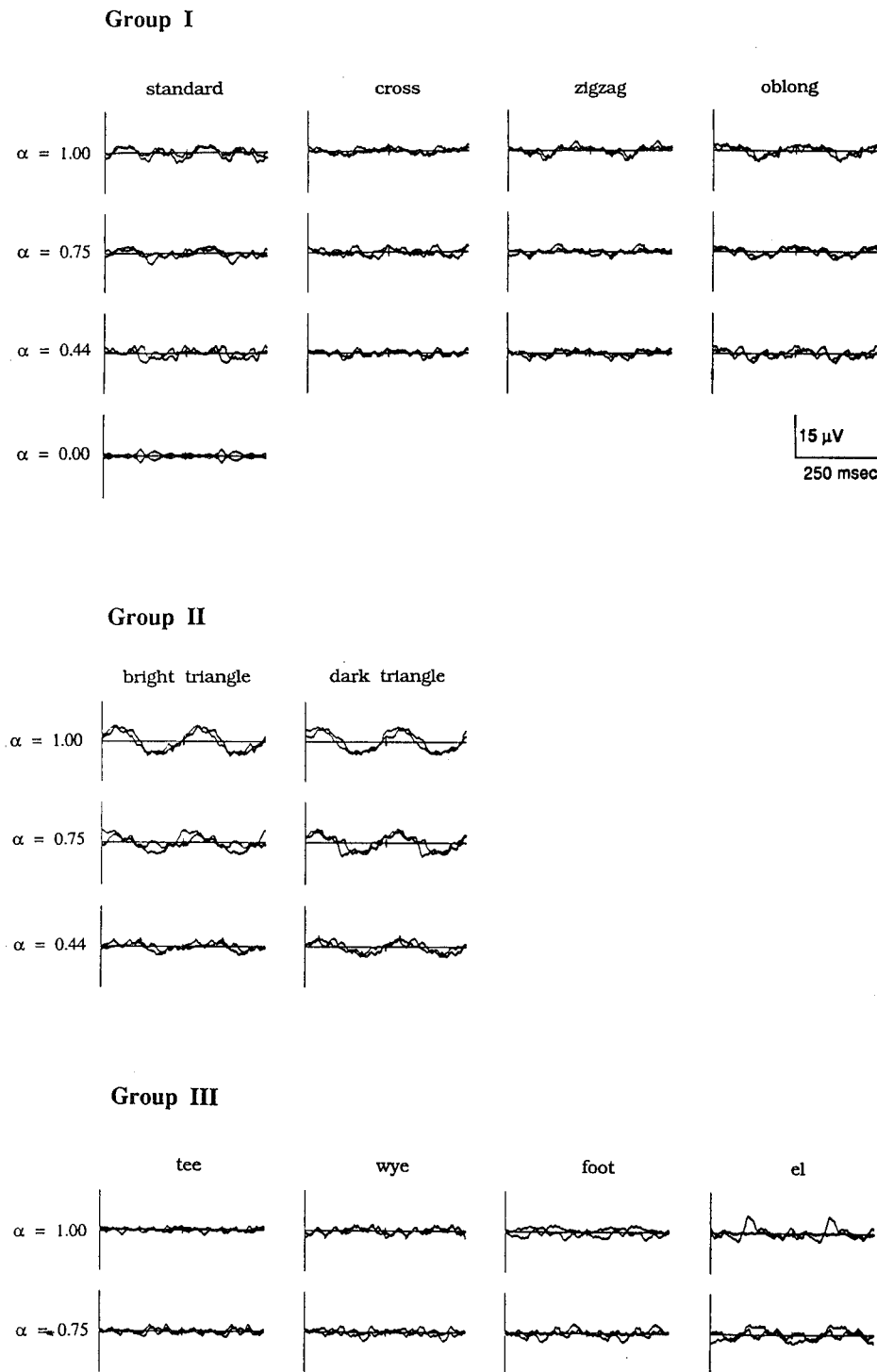


Fig. 5(C)

Fig. 5. Symmetric and antisymmetric components of the VEP elicited by interchange between structured isodipole textures and the random texture. Part A: symmetric component. Part B: antisymmetric component. Part C: pattern-dependent symmetric subcomponent. Waveforms are derived from the data of Fig. 4. Responses to interchanges of examples of the random texture are illustrated as the " $\alpha = 0$ " response to the standard texture in Group I.

correlation and response amplitude depended strongly on the texture type. In Table 3, we have summarized the dependence of the antisymmetric VEP component on local correlation α and texture type G by the slope of the regression of first-harmonic amplitude on α , which we will denote $K(G)$. In general, the assumption of a strict proportionality between α and response amplitude, coupled with the hypothesis of constant response phase, provided fits to the observed first-harmonic amplitudes with a mean-squared error of $0.3 \mu\text{V}$ (range $0.1\text{--}0.9 \mu\text{V}$ across all 10 textures and six subjects). For comparison, the run-to-run reproducibility of the antisymmetric response itself was typically $0.5 \mu\text{V}$.

As seen in Table 3, the values of $K(G)$ found for Groups I and II textures averaged about three times those found for Group III textures. This is in accord with the psychophysical findings as summarized in Table 2, although the difference is somewhat more

marked in the psychophysical responses. There is no consistent difference between Groups I and II textures, which corresponds to the substantial overlap of performance psychophysically.

Rank-order of values of $K(G)$ within groups corresponds to the psychophysical data of Table 2 in a general fashion. Within Group I, largest values of $K(G)$ are found typically for the standard texture; smallest values are found typically for the cross texture. Within Group III, largest values of $K(G)$ are found typically for the tee and foot textures, and smallest values are found typically for the wye texture.

However, there are also some consistent differences between the VEP and psychophysical results. For example, the "oblong" texture elicited the second-largest antisymmetric VEP response among Group I textures, but it was the hardest to detect psychophysically [smallest $R(G)$]. Other than the superiority of

Table 3. The dependence of the antisymmetric VEP component on local correlation α and texture type G . Response amplitudes are summarized by the best-fitting proportionality $K(G)$ between the size of the first harmonic and the local correlation α , with units of $\mu\text{V}/\text{unit correlation}$. Response phases are in units of π rad. The geometric mean is used for $K(G)$ and the arithmetic mean is used for phase

Texture type		Subject						Mean
		JV	MC	SU	GG	CI	RB	
Group I								
Standard	$K(G)$	6.15	4.90	4.39	3.05	5.23	3.91	4.50
	phase	1.16	1.13	1.10	1.04	1.06	1.10	1.10
Cross	$K(G)$	3.37	2.26	2.17	1.52	2.76	1.88	2.25
	phase	1.08	0.85	1.02	0.97	0.98	1.04	0.99
Zigzag	$K(G)$	3.96	2.72	2.28	2.05	2.71	2.37	2.61
	phase	1.18	0.92	1.08	0.98	1.07	1.10	1.05
Oblong	$K(G)$	4.92	2.64	3.07	1.94	2.80	2.25	2.80
	phase	1.11	0.95	1.08	0.98	1.06	1.08	1.04
Group mean	$K(G)$	4.48	2.99	2.86	2.07	3.23	2.50	2.93
	phase	1.13	0.96	1.07	0.99	1.04	1.08	1.05
Group II								
Bright triangle	$K(G)$	4.35	2.86	3.34	1.81	2.64	1.72	2.64
	phase	1.04	0.90	0.99	0.85	0.93	1.00	0.95
Dark triangle	$K(G)$	3.99	3.06	3.72	2.47	3.80	2.53	3.20
	phase	1.14	0.79	0.95	0.82	0.93	1.01	0.94
Group mean	$K(G)$	4.17	2.96	3.52	2.11	3.17	2.09	2.91
	phase	1.09	0.80	0.97	0.84	0.93	1.01	0.94
Group III								
Tee	$K(G)$	1.21	0.96	1.54	1.51	0.84	1.10	1.16
	phase	0.87	0.57	0.84	0.74	0.84	0.89	0.79
Wye	$K(G)$	0.51	0.58	1.38	0.71	0.48	0.53	0.64
	phase	0.83	0.92	0.80	0.88	0.47	0.88	0.80
Foot	$K(G)$	1.19	1.21	1.03	0.74	0.89	0.41	0.86
	phase	1.11	0.58	0.82	0.74	1.00	0.93	0.86
El	$K(G)$	1.08	0.71	1.04	0.85	0.60	0.25	0.68
	phase	0.80	0.40	0.79	0.74	0.94	1.17	0.81
Group mean	$K(G)$	0.94	0.83	1.22	0.91	0.68	0.49	0.81
	phase	0.90	0.62	0.82	0.78	0.82	0.97	0.82

the standard texture to the other three by both VEP and psychophysical criteria, there was no consistent correlation between the rank-order of VEP and psychophysical responses to Group I textures.

The VEP data provide information on anti-symmetric response dynamics (phase) in addition to overall size [$K(G)$]. Examination of phase (Table 3) reveals a consistent difference between responses to the three groups of textures. Average response phase was 1.05π radians for Group I textures, 0.94π radians for Group II textures, and 0.82π radians for Group III textures. Although there was some variation between subjects, there was little or no overlap of phases between groups in any given subject. These differences of 0.11 – 0.23π radians within subjects but between groups are large in comparison with the reproducibility of response phases measured on separate occasions (e.g. this experiment and the "propagated decorrelation" experiment, or the data of Victor & Conte, 1989), which was typically 0.04π rad. The effective latency difference for 2.12 Hz response corresponds to a 25 msec lag for Group II textures compared to Group I textures, and a 54 msec lag for Group III textures compared to Group I textures. Differences in dynamics of responses to the various textures will be further investigated in the next section.

VEP responses: principal components analysis. Inspection of the antisymmetric component and the pattern-dependent symmetric subcomponent of the VEP suggested that responses to distinct isodipole textures differed in dynamics, as well as in overall size. In order to determine whether the dependence of VEP dynamics on pattern type was systematic, we pursued a principal components analysis (Cattell, 1978) of the pattern-dependent VEP waveforms. The pattern-dependent VEP response to the texture $T_G(1, 1)$ is the sum of the antisymmetric component (e.g. Fig. 5B) and the pattern-dependent symmetric subcomponent (e.g. Fig. 5C). To begin the principal components analysis, we represent this response $\mathbf{R}(G)$ by a vector of its Fourier components. In order to improve signal-to-noise, we consider only the first eight Fourier components (Gutowitz, Zemon, Victor & Knight, 1986). Thus:

$$\mathbf{R}(G) = [R_1(G), R_2(G), \dots, R_8(G)] \quad (11)$$

where $R_n(G)$ is a complex number which represents the amplitude and phase of the n th

Fourier component of the pattern-dependent response $\mathbf{R}(G)$.

We now attempt to represent the responses $\mathbf{R}(G)$ as linear combinations of a small number H of "principal components" \mathbf{P}_h :

$$\mathbf{R}(G) = \sum_{h=1}^H w_h(G) \mathbf{P}_h. \quad (12)$$

The weights w_h are normalized by:

$$\sum_G [w_h(G)]^2 = 1. \quad (13)$$

If the representation (12) succeeds for $H = 1$, then all pattern-dependent responses $\mathbf{R}(G)$ can be regarded as scalar multiples of a common response waveform \mathbf{P}_1 . However, if there is a systematic dependence of dynamics on pattern type, then at least two principal components \mathbf{P}_1 and \mathbf{P}_2 will be required for the representation (12). The weights $w_1(G)$ and $w_2(G)$ will then indicate the contribution of each principal component \mathbf{P}_1 and \mathbf{P}_2 to the pattern-dependent response $\mathbf{R}(G)$.

In general, we found that two principal components ($H = 2$) were required for a good fit to the data. A test of the significance of the second principal component is its consistency across subjects, which we will examine below. Unfortunately, even with H fixed at the value 2, the representation (12) is not unique (Cattell, 1978). The representation (12) determines the H -dimensional subspace which most nearly contains (in a least-square sense) the responses $\mathbf{R}(G)$. This best-fitting subspace is unique. However, the choice of basis vectors $w_h(G)$ within that subspace is not uniquely determined. A change of basis vectors $w_h(G)$ induces a linear transformation among the principal components \mathbf{P}_h . Thus, the particular values $w_h(G)$ and the particular waveforms \mathbf{P}_h cannot be regarded as having a direct interpretation. For this reason, we do not further analyze the individual values $w_h(G)$, but instead we analyze algebraic "invariants" constructed from them (Gutowitz et al., 1986). These invariants $I(G, G')$ are functions of two texture types (G and G'), and are defined by:

$$I(G, G') = w_1(G)w_2(G') - w_2(G)w_1(G'). \quad (14)$$

The determinant-like nature of $I(G, G')$ guarantees that its value is independent of a change in basis vectors, up to a common scale factor (Gutowitz et al., 1986). The invariants have an intuitive interpretation: its magnitude expresses

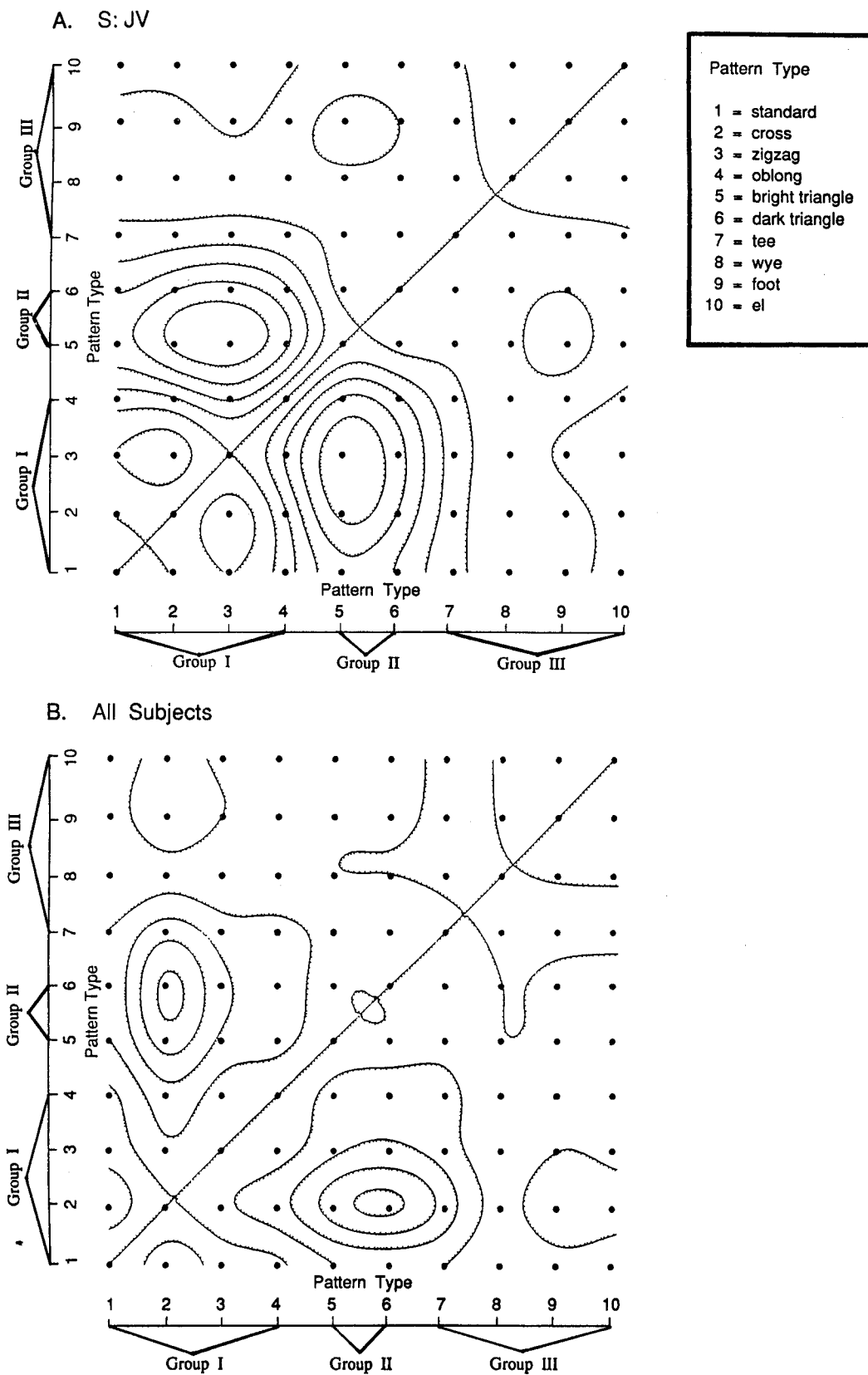


Fig. 6. Contour maps of the invariants $I(G, G')$. The height of the surface corresponding to a pair of textures G and G' indicates the extent to which they represent distinct linear combinations of the first two principal components. Each contour line represents 0.1 unit, and tickmarks point downhill. Part A: invariants calculated from data of Fig. 5. Part B: average of invariants across all subjects.

the extent to which the responses to textures G and G' fail to be scalar multiples of each other.

In Fig. 6A, we show a contour map of the invariant $I(G, G')$ derived from the pattern-dependent VEP components of Fig. 5. The diagonal symmetry of the plot is due to the identity $I(G, G') = -I(G', G)$. The largest absolute value of $I(G, G')$ occurs when G is the bright triangle texture and G' is the zigzag texture. $I(G, G')$ is essentially zero when either one of the two textures is from Group III. This is simply a consequence of the fact that these textures produce only very small pattern-dependent responses, and consequently their weights w_h are close to zero.

A similar set of invariants was constructed for the other five subjects. To avoid biasing the average in favor of those subjects whose overall VEP amplitudes were large, data from individual subjects were weighted by the reciprocal of the overall pattern-dependent amplitude. Figure 6B shows the cross-subject average of the invariants. The major features seen in the individual subject (Fig. 6A) again appear. The largest values of the invariant are seen when one of the textures is a Group I texture and the other is a Group II texture, and very small values are found when either of the texture is a Group III texture. Since the invariant $I(G, G')$ indicates the difference between the dynamics of responses to textures G and G' , we conclude that there are significant differences in dynamics between Groups I and II textures. As seen from the height of the invariant plots, these differences are two to three times as large as differences within textures of Group I, and four or more times as large as differences within textures of Group II.

We now turn to an examination of the consistency of the first two principal components across subjects. For any pair of subjects s_1 and s_2 , we begin with the representations in terms of principal components:

$$\begin{aligned} \mathbf{R}_{s_1}(G) &= \sum_{h=1}^H w_{h,s_1}(G) \mathbf{P}_{h,s_1} \\ \mathbf{R}_{s_2}(G) &= \sum_{h=1}^H w_{h,s_2}(G) \mathbf{P}_{h,s_2}. \end{aligned} \quad (15)$$

We form the dot-product of weights from each pair of principal components, $D_{h_1, h_2}(s_1, s_2)$:

$$D_{h_1, h_2}(s_1, s_2) = \sum_G w_{h_1, s_1}(G) \cdot w_{h_2, s_2}(G). \quad (16)$$

For any pair of subjects, the matrix of dot-products D_{h_1, h_2} represents the extent to which the weights $w_{h_1, s_1}(G)$ for principal component h_1 from subject s_1 correspond to the weights $w_{h_2, s_2}(G)$ for principal component h_2 from subject s_2 . If the weights correspond exactly, then $D_{h_1, h_2} = 1$. If there is no correlation among the weights, then $D_{h_1, h_2} = 0$.

As pointed out above, the particular values of the weights w_{h, s_1} depend on the (arbitrary) choice of linear combinations for the principal components $P_h(G)$. Thus, for two or more principal components, values of the individual weights (and hence of D_{h_1, h_2}) cannot be interpreted directly. However, we again circumvent this problem by the construction of an invariant which is independent of linear transformations among the two principal components. This invariant is the determinant of the matrix of dot-products $D_{h_1, h_2}(s_1, s_2)$. It will be denoted $X^{(2)}(s_1, s_2)$, and is given by:

$$\begin{aligned} X^{(2)}(s_1, s_2) &= D_{1,1}(s_1, s_2) \cdot D_{2,2}(s_1, s_2) \\ &\quad - D_{1,2}(s_1, s_2) \cdot D_{2,1}(s_1, s_2). \end{aligned} \quad (17)$$

Analogous cross-invariants $X^{(n)}(s_1, s_2)$ may be constructed from any number n of principal components extracted from two subjects, as the determinant of the matrix of dot-products $D_{h_1, h_2}(s_1, s_2)$. For a single principal component, the corresponding invariant $X^{(1)}(s_1, s_2)$ is simply the dot-product $D_{1,1}(s_1, s_2)$ itself.

This procedure has a geometric interpretation. A set of weights $w_h(G)$ may be regarded as a unit vector in a response space, which expresses how heavily principal component h contributes to responses to each of the textures. Let us first consider the case in which there is only one significant principal component [$H = 1$ in equation (12)]. If the weights corresponding to the first principal components from two subjects s_1 and s_2 are similar, then the vector $w_{1, s_1}(G)$ and the vector $w_{1, s_2}(G)$ will point in similar directions. The quantity $X^{(1)}(s_1, s_2) = D_{1,1}(s_1, s_2)$ is the cosine of the angle between the weight vector $w_{1, s_1}(G)$ and the weight vector $w_{1, s_2}(G)$. It thus expresses the similarity between the first principal components in subjects s_1 and s_2 .

When there is more than one significant principal component, the situation is more complex. The representation (12) identifies an H -dimensional subspace which contains the responses $\mathbf{R}(G)$. The invariant (17) is the cosine of the dihedral angle between the response

Table 4. Invariants $X^{(1)}(s_1, s_2)$ and $X^{(2)}(s_1, s_2)$ for all pairs of subjects (s_1, s_2). Values of $X^{(n)}$ close to 1 indicate close correspondence between the weights of the n -term principal components representations (15) for the two subjects. Significant correspondences at $P = 0.05$ are indicated by *, and at $P = 0.01$ are indicated by **

Subject s_2		JV	MC	Subject s_1			
				SU	GG	CI	RB
MC	$X^{(1)}$	0.987**					
	$X^{(2)}$	0.462					
SU	$X^{(1)}$	0.954**	0.946**				
	$X^{(2)}$	0.640*	0.835**				
GG	$X^{(1)}$	0.934**	0.922*	0.981**			
	$X^{(2)}$	0.429	0.716*	0.794**			
CI	$X^{(1)}$	0.970**	0.984**	0.941**	0.943**		
	$X^{(2)}$	0.624*	0.881**	0.841**	0.769**		
RB	$X^{(1)}$	0.974**	0.988**	0.957**	0.953**	0.994**	
	$X^{(2)}$	0.569	0.653*	0.830**	0.752**	0.837**	

subspaces for subjects s_1 and s_2 . Thus, it is equal to 1 when the subspaces coincide, and equal to zero when they are perpendicular. We note that although the subjects' response subspaces are uniquely determined by the representation (12), the particular weights are not uniquely determined. However, the value of the invariant (17) is independent of the choice of weights, since it is an intrinsic property of the subspaces themselves.

Values of the invariants $X^{(1)}(s_1, s_2)$ and $X^{(2)}(s_1, s_2)$ for all pairs of subjects are displayed in Table 4. The invariants $X^{(1)}$ range from 0.922 to 0.994. Values close to 1 correspond to close correspondence across subjects; values close to 0 indicate a lack of correspondence. The statistical significance of the calculated values for $X^{(1)}$ can be determined from the distribution of values that would be obtained for random weight vectors w_{1,s_1} and w_{1,s_2} . By this measure, all 15 values are significant at $P = 0.01$ (critical value for $X^{(1)}$ is 0.701).

Values of the invariants $X^{(2)}$ range from 0.429 to 0.881. We consider two measures of statistical significance. The naive measure is derived from the distribution of values of $X^{(2)}$ that would be obtained from random weight vectors w_{1,s_1} , w_{2,s_1} , w_{1,s_2} , and w_{2,s_2} (subject to the condition that weights for each subject are orthonormal). By this measure, all 15 values of $X^{(2)}$ are significant at $P = 0.01$ (critical value for $X^{(2)}$ is 0.420). This indicates that the response planes for any pair of subjects are more aligned than would be expected by chance alone. However, it does not really indicate that the second principal component is significant across subjects, since the nonrandom alignment of response planes may simply be due to the alignment of the first principal component. To address the signifi-

cance of the second component *per se*, we calculated another distribution of values of $X^{(2)}$. This distribution assumed that the first pair of weights w_{1,s_1} and w_{1,s_2} are equal, and the second pair w_{2,s_1} w_{2,s_2} are random. It thus allows for testing of the incremental significance of the second principal component. By this measure, eight of the 15 values of $X^{(2)}$ are significant at $P = 0.01$ (critical value for $X^{(2)}$ is 0.730), and 12 of the 15 values of $X^{(2)}$ are significant at $P = 0.05$ (critical value for $X^{(2)}$ is 0.616). This more stringent criterion is the one used for indicating significance in Table 4, and will be the criterion used subsequently.

In an analogous fashion, cross-subject consistency between the first three principal components would be manifest by a non-random distribution of the third invariant $X^{(3)}$, which is defined as a 3×3 -determinant similar to equation (17). To assess incremental significance of the third principal component, we calculated a distribution of values of $X^{(3)}$ assuming that the first two pairs of weights w_{h,s_1} and w_{h,s_2} ($h = 1$ and 2) determine the same plane, and the third pair w_{3,s_1} w_{3,s_2} are random. This analysis revealed no evidence of significance of the third invariant $X^{(3)}$: one value out of 15 was significant at $P = 0.05$, with a critical value for $X^{(3)}$ of 0.651.

The analysis described above was also performed for the antisymmetric component alone (e.g. Fig. 5B), and the pattern-dependent symmetric subcomponent alone (e.g. Fig. 5C). For the antisymmetric component, cross-subject invariants $X^{(1)}$ ranged from 0.920 to 0.990, and all were significant at $P = 0.01$. The cross-subject invariants $X^{(2)}$ ranged from 0.271 to 0.866. By the more stringent test of significance described above, 6 of the 15 values were

significant at $P = 0.01$ and 10 were significant at $P = 0.05$.

For the symmetric subcomponent, cross-subject invariants $X^{[1]}$ ranged from 0.028 to 0.944. Eight of the 15 values were significant at $P = 0.01$ and twelve were significant at $P = 0.05$. The cross-subject invariants $X^{[2]}$ ranged from 0.022 to 0.688. By the more stringent test of significance, none of the 15 values were significant at $P = 0.01$ and only two were significant at $P = 0.05$.

Thus, we conclude that for these textures, the pattern-sensitive VEP's are described by a two-parameter family of responses (12). Across subjects, the near-unity values of the invariants $X^{[1]}$ and $X^{[2]}$ (Table 4) indicate a similarity in weights w_1 and w_2 with which the two principal components combine to produce a response. Both principal components are manifest in the antisymmetric component of the response, but only one principal component is present in the symmetric subcomponent. In a qualitative sense, the requirement for *two* principal components, rather than a single component, is the difference between responses to Groups I and II textures. This is seen from the peak of the invariant $I(G, G')$ when one texture is from Group I and the other is from Group II (Fig. 6).

VEP responses: propagated and sporadic decorrelation. The results described so far were obtained from fully-correlated textures $T_G(1, 1)$, or from textures with sporadic decorrelation $T_G(1, \alpha)$. We now compare these responses to VEPs elicited by textures $T_G(\alpha, 1)$ with propagated decorrelation. The purpose of this comparison is to estimate the spatial scale of the interactions that generate the pattern-sensitive response.

The estimate is based on the relationship between the correlation properties of the textures $T_G(1, \alpha)$ and $T_G(\alpha, 1)$. Both textures have the same degree of local correlation α . Textures $T_G(1, \alpha)$ with sporadic decorrelation only ($c_{\text{prop}} = 1, c_{\text{spor}} = \alpha < 1$) have preserved long-range correlation structure. Textures $T_G(\alpha, 1)$ with propagated decorrelation only ($c_{\text{prop}} = \alpha \lesssim 1, c_{\text{spor}} = 1$) have no long-range correlation structure. [These statements follow from equation (10), with $n(S) = 4$ and $n(S/G)$ large.] If the spatial interactions underlying VEP generation are all short-range, then responses to the textures should depend only on the local correlation α , and not on whether long-range correlations are present. Conversely, if the interactions underlying VEP generation are long-

range, then responses to a texture $T_G(1, \alpha)$ with preserved long-range correlation should be substantially larger than the responses to a texture $T_G(\alpha, 1)$ whose long-range correlations diminish rapidly with distance. As described in detail in Victor and Conte (1989), a quantitative comparison of responses obtained with sporadic and propagated decorrelation leads to an estimate of the distribution and overall spatial scale D of nonlinear interactions which generate the VEP.

For three subjects (JV, MC, RB) and three textures (standard, oblong, and bright triangles), we measured antisymmetric VEP components at five levels of sporadic and propagated decorrelation. In all cases, as local correlation α decreased to zero, responses decreased more rapidly for textures $T_G(\alpha, 1)$ with propagated decorrelation than for textures $T_G(1, \alpha)$ with sporadic decorrelation. We applied the method of Victor and Conte (1989) to interpret this finding in terms of the form and spatial scale of long-range interactions. For all textures, a one-dimensional exponential weighting of spatial interactions provided a better fit than did two-dimensional exponential weightings, or one- and two-dimensional Gaussian weightings, or uniform weightings.

The spatial scale D of the best-fitting weightings is shown in Table 5. Although a threefold range of spatial scales D is found across subjects and across texture types, there are no striking trends. The variation of estimated scale D across subjects is of a similar magnitude as the variation across texture types within a subject. This variability does not represent experimental uncertainty: estimates of spatial scale recorded on separate occasions over one year apart agreed within 8% (values in parentheses in Table 5).

DISCUSSION

We begin by summarizing the present experimental results in the context of our previous

Table 5. The spatial scale D (min) of interactions for three isodipole textures. D is calculated from the comparison of responses to textures $T_G(1, \alpha)$ with sporadic decorrelation with responses to textures $T_G(\alpha, 1)$ with propagated decorrelation, with 4 min checks. Values in parentheses were obtained in a previous study (Victor & Conte, 1989), and are included to illustrate reproducibility

Texture type	Subject		
	JV	MC	RB
Standard	8.1 (7.5)	13.4 (12.4)	13.9
Oblong	11.5	19.5	12.9
Bright triangle	14.5	22.5	7.9

studies. On qualitative grounds, this will permit us to exclude a variety of models for our findings, and will suggest an architecture for models which might succeed. We then present one simple example of such a model, and discuss its salient features.

Summary of previous results

In previous studies (Victor, 1985; Victor & Zemon, 1985), we showed that alternation between two isodipole textures: the even texture [here, $T_{\text{standard}}(1, 1)$] and the odd texture [here, $T_{\text{standard}}(-1, 1)$] generated a robust VEP. The statistics of the isodipole textures implied that this VEP and the salient perceptual differences between these textures were generated by a highly nonlinear neural mechanism. Such a mechanism must have a formal order at least four, and must include nonlinear interactions between four distinct points in space. In principle, summation of neural signals followed by a nonlinearity of high order (such as rectification) possesses the requisite complexity. However, quantitative analysis excluded this relatively simple mechanism (Victor 1985), and, by implication, mechanisms composed of linear superpositions of such subunits.

In later studies (Victor & Conte, 1989), we explored VEP and psychophysical responses to the standard texture constructed with a range of check sizes and degrees of local and global correlation structure. This allowed us to test additional models within the linear-nonlinear framework, in which the linear stage was more elaborate than simple spatial pooling. Models with these more sophisticated linear stages (such as an elongated bar, edge-detector, or center-surround profile) could not account for several qualitative features of the responses to the standard even [$T_{\text{standard}}(1, 1)$] and odd [$T_{\text{standard}}(-1, 1)$] textures. However, this analysis did suggest a possible framework for a successful model: an array of sufficiently complex nonlinear subunits over a wide area, whose outputs were combined in a manner which was exponentially decreasing with distance. At any position in visual space, the spatial scales of these subunits appeared to vary over a 4–8-fold range.

Summary of current results

In this study, we examined VEP and psychophysical responses to a variety of texture classes, which shared the isodipole property but differed in the choice of recursion relation, or “glider” (Table 1). This allowed us to explore which local

features of the stimulus were responsible for the perception of pattern and the production of a pattern-dependent VEP response. The main qualitative finding is that textures based on some gliders (Groups I and II) led to rapid psychophysical detection of pattern (Fig. 3, Table 2) and a large pattern-dependent VEP (Figs 4 and 5, Table 3). Textures based on other gliders (Group III) were more difficult to detect psychophysically and produced much smaller pattern-dependent VEPs. We divided the textures that were readily distinguishable into two groups based on the size of the glider (Group I—4 checks, Group II—3 checks).

Although psychophysical detection and VEP amplitudes were similar in both groups, two lines of evidence pointed to physiological differences between responses to these groups. There was a consistent difference in VEP phase, with Group I responses leading Group II responses by about a twentieth of a cycle (Table 3), corresponding to a latency shift of 25 msec. Principal components analysis showed that VEP responses could be adequately reconstructed as a sum of two waveforms. The requirement for two waveforms was primarily due to differences between Group I responses and Group II responses (Table 4).

Other components of the analysis failed to reveal as clear a pattern of consistency across subjects. These included the rank-order of psychophysical and VEP responses within Group I and Group II (Tables 2 and 3), and the spatial scale of interactions as determined from VEP responses to textures with propagated decorrelation (Table 5).

Relationship of VEP and psychophysical studies

We have pursued the psychophysical studies in parallel with the VEP studies in order to be sure that the neural interactions measured in the VEPs are relevant to pattern vision. However, psychophysical studies do not provide for as detailed and direct a measure of response dynamics as do the VEP studies (e.g. Tables 3 and 4). Additionally, the VEPs provide a bridge to studies that can only be carried out in the experimental animal. In preliminary studies, (Purpura & Victor, unpublished results) isodipole VEPs which share the main features of the responses presented here have been recorded epicortically over V1 of the macaque. Thus, in modeling the neural circuitry underlying texture responses, we should focus on processes likely to be present in striate cortex.

Implications for cortical processing

We now turn to consideration of models for early visual processing that account for these findings. The major result that we seek to account for is that textures based on some gliders lead to immediate perceptual salience and a large VEP, and others do not. Although a detailed analysis reveals subtle differences between Groups I and II responses, the VEP responses and psychometric curves are similar (Figs 3–5). For this reason, we feel justified in attempting to account for the perceptual salience of textures in both classes with a single model, although we will later discuss possible mechanisms underlying the more subtle differences between Groups I and II responses.

As discussed above, a successful model to account for discrimination of these isodipole textures must be highly nonlinear. In Victor and Conte (1989), we suggested a framework consisting of an array of sufficiently complex local nonlinearities, whose outputs are combined by a second nonlinear stage. As a preliminary step, we will consider potentially plausible alternative kinds of models whose structure is unrelated to this framework. These models may be excluded on the basis of qualitative features of failure to separate the readily-distinguishable textures (Groups I and II) from those that were not readily-distinguishable (Group III). Then, we turn to a simple, physiologically-plausible example of the two-nonlinearity model, and discuss how our data permit parameters of the model to be refined.

A role for information content?

Perception of a visual pattern implies perception of redundancy; the existence of a pattern implies that one portion of the visual texture places constraints on a second portion. Thus, textures in which structure is evident may be thought of as having a lower information content (i.e. greater redundancy) as perceived by the visual system. To what extent is the visually-perceived information content simply related to the usual formal notion (Shannon, 1948) of information content, as applied to the isodipole textures?

The Shannon information content of a region R of a visual texture T is given by a sum over all possible allowed configurations λ of the texture:

$$I[R, T] = -\sum_{\lambda} p(\lambda) \log p(\lambda) \quad (18)$$

where “log” is the logarithm to the base 2. An example will clarify the relationship of low information content to pattern. For the random texture $T(0, 0)$, all of the 2^{ij} ways of coloring the checks of an $i \times j$ region are equally likely. Thus, each texture configuration has probability 2^{-ij} , and $I[R, T(0, 0)] = ij \log 2$. In contrast, for the fully-correlated even texture $T_{\text{standard}}(1, 1)$, the configuration of an $i \times j$ region is determined by the configuration of the first row and the first column. Thus, there are only 2^{i+j-1} configurations, each of which are equally likely. It follows that

$$I[R, T_{\text{standard}}(1, 1)] = (i + j - 1) \log 2.$$

Thus, the information content of a patch of a random texture is proportional to the number of checks it contains, while the information content of a patch of the fully-correlated standard texture is proportional to the number of checks in its first row and first column.

For any glider G , the configuration of a region of a fully-correlated texture $T_G(1, 1)$ is completely determined by the random assignment of check polarities to one or more initial rows and columns. After assignment to a sufficiently large “initial set” (Gilbert, 1980) the interior of the texture is determined. Thus, the information content of a region of the texture is proportional to the minimum number of checks that must be specified. Table 6 lists how many rows and columns must be specified to determine textures based on each kind of glider. As seen from Table 6, the hypothesis that textures with greater redundancy have more salient structure than textures with less redundancy is not supported. The least-redundant texture (oblong) has salient structure, while two of the most redundant textures (tee and el) have relatively inapparent structure.

A role for symmetry?

The information-theoretic analysis shows that not all spatial correlations are equally relevant to the visual perception of structure, or redundancy. Symmetry is a global feature readily recognized by the visual system. We considered the possibility that symmetry of the glider, and hence of the texture statistics, is the crucial characteristic which separates the textures with salient structure from those without salient structure. Symmetry properties of the gliders are listed in Table 6. Two of the textures with the most salient structure are

Table 6. Properties of selected isodipole textures and their gliders. The model responses were calculated from the model of Fig. 7, with parameters $N = 6$, $h = 0.6$, $r = 0.5$, $r_a = 0.5$, $r_b = 0.5$, $\theta = 90$ deg, as described in the text. The bright and dark triangle textures, each constructed from the same triangle glider, share all properties listed

Texture type	Order of glider	Sum of squared distances	Maximum squared distance	Symmetry	Information content	Response of proposed model
Group I						
Standard	4	8	2	8-fold	1 row, 1 col	0.0348
Cross	4	16	4	8-fold	2 rows	0.0292
Zigzag	4	12	5	2-fold	1 row, 1 col	0.0311
Oblong	4	20	5	4-fold	1 row, 2 cols	0.0297
Group II						
Triangle	3	4	2	2-fold	1 row	0.0289
Group III						
Tee	4	11	4	2-fold	1 row	0.0242
Wye	4	16	5	2-fold	2 rows	0.0264
Foot	4	15	5	1-fold	1 row, 1 col	0.0279
El	4	14	5	1-fold	1 row	0.0253

generated by gliders with the 8-fold symmetry of the square (standard and cross), while two of the textures without salient structure are generated by gliders without symmetry (foot and el). However, examination of the remaining gliders shows that symmetry *per se* is not the crucial feature which determines the percept of structure. Some gliders with two-fold symmetry generate textures with readily-apparent structure (oblong and triangle), while others do not (tee and wye).

We also note that it is possible to manipulate symmetry by changing the lattice upon which the texture is constructed. For example, the square lattice may be replaced with a hexagonal one. This reduces the symmetry of the standard and cross gliders from 8 to 4-fold, while it increases the symmetry of the wye glider from 2 to 6-fold. Informal observations show that the qualitative percept of presence or absence of visual structure is not changed by this manipulation.

A role for glider size?

Let us consider the possibility that the overall size of the glider (which provides the basic constraint that defined the texture) determines whether a statistical constraint is detected visually. Two measures of the size of a glider are listed in Table 6: the sum of the squares of the distances among all cells of the glider, and the square of the maximum distance between two cells of the glider. As seen in Table 6, neither measure separates the textures with salient structure from the textures without salient structure.

Density of correlations in glider iterates

Since neither glider symmetry nor size determine which textures are visually salient, we are compelled to examine more subtle properties of the textures. Since the textures are generated by repeatedly applying a recursion rule specified by a glider, we turn to properties of these glider iterates. Consider the standard texture. Its defining rule constrains the number of bright checks in a 2×2 region to be even. However, since the number of bright checks in an overlapping 2×2 -region must also be even, it follows that the number of bright checks at the corners of a 2×3 -region is also even. Thus, iteration of the local recursion rule induces a relationship between the states of checks at distances greater than the extent of the glider. For the even texture, appropriate iteration of the recursion rule shows that the number of bright checks at the corners of an $(I + 1) \times (J + 1)$ rectangle is always even.

This geometric property corresponds to an algebraic property of generating functions of regions and gliders. As discussed in Methods, the states of checks in a set S are correlated if (and only if) the generating function $\hat{S}(x, y)$ (equation 7) corresponding to S is divisible (mod 2) by the generating function $\hat{G}(x, y)$ corresponding to the glider G . Thus, the four-membered sets S whose checks are correlated corresponds to the four-term polynomials $\hat{S}(x, y)$ which are divisible by $\hat{G}(x, y)$. As seen from the example following equation (8), for the standard texture, all polynomials of the form $\hat{S}(x, y) = (1 + x^I)(1 + y^J)$ indeed have $\hat{G}_{\text{standard}}(x, y)$ as a factor.

Thus, for the standard texture, there are many four-membered sets of checks substantially larger than the glider whose states are correlated. More precisely, the number of fourth-order correlations within a distance D of a particular check is proportional to D^2 . This property is common to the other Group I textures as well, although the proportionality constant varies from texture to texture. One way to see this is to note that correlated sets correspond to polynomials which are divisible by the glider's generating function (equation 10), and the generating functions of the other Group I textures are algebraically related to $\hat{G}_{\text{standard}}$:

$$\begin{aligned} \hat{G}_{\text{cross}}(x, y) &= x + y + x^2y + xy^2 \\ &= y\hat{G}_{\text{standard}}(xy, x/y) \end{aligned} \quad (19)$$

$$\begin{aligned} \hat{G}_{\text{zigzag}}(x, y) &= 1 + y + xy + xy^2 \\ &= \hat{G}_{\text{standard}}(xy, y) \end{aligned} \quad (20)$$

$$\begin{aligned} \hat{G}_{\text{oblong}}(x, y) &= 1 + x + y^2 + xy^2 \\ &= \hat{G}_{\text{standard}}(x, y^2). \end{aligned} \quad (21)$$

On the other hand, Group III textures have far fewer 4-element sets which are correlated. For these textures, the asymptotic number of sets of size D or less appears to be $\log D$ (tee, wye, or el) or $\log D^2$ (foot). The logarithmic dependence follows from the fact that for all four-element gliders, polynomials of the form $\hat{G}^{2^m}(x, y)$ contain only four nonzero terms (mod 2), and all are necessarily divisible by \hat{G} .

For textures in Group II, the asymptotic number of such sets appears to be greater than that of any of the Group III textures. This may be seen from the fact that all polynomials of the form $[\hat{G}_{\text{triangle}}(x, y)]^{2^m}$ have three elements, so that superpositions of the corresponding sets and their translates (with single checks in common) produce many 4-element correlated sets of size D .

The difficulty in making these statements rigorous lies not in providing lists of 4-element sets which are correlated, but in proving that the lists are exhaustive. For the tee, foot, and el textures, the asymptotic estimates above may be rigorously proven by tedious counting arguments; for the other textures, no such proof is as yet available.

A candidate model

We draw two conclusions from the above analysis: (i) the visual system only detects pat-

tern in textures with a high degree of mathematical redundancy if the spatial organization of the correlation structure is appropriate; and (ii) a model which distinguishes a salient correlation structure from one which is much less visually salient must exploit correlations over a distance substantially greater than the glider itself. Previously (Victor & Conte, 1989), we had concluded that models which contain a local nonlinearity followed by a second nonlinearity of wider spatial extent might account for responses to the standard texture. We now present an example of such a model, and show that it accounts for our current experimental results.

Figure 7 illustrates the spatial organization of a computational element of the proposed model. Signal processing begins with linear spatial summation within a circular Gaussian profile of radius r . That is, the output of this component of the model is determined by convolving the sensitivity function

$$S(x, y) = \frac{1}{\pi r^2} \exp[-(x^2 + y^2)/r^2] \quad (22)$$

with the stimulus luminance. In this and all other discussion of the model, length parameters (such as r) are in units of check size, and stimulus contrast is taken to be unity. Note that the Gaussian profile (22) is normalized to unit volume, so that its response to a stimulus of unit contrast must lie between -1 and 1 .

We postulate that the output of the Gaussian profiles are combined in antagonistic pairs in a linear fashion. As shown in Fig. 7, the centers of the Gaussians of each pair are separated from each other by a distance $2r_a$. The result of this summation is half-wave rectified. Thus, each pair of linearly-summing regions constitutes a nonlinear subunit (Hochstein & Shapley 1976), which is a primitive edge detector whose output ranges from 0 to 2. The maximum output occurs when one profile of the pair is entirely within a bright region, and the other profile is entirely within a dark region.

These subunits are positioned along a line, separated from each other by a distance r_b , and oriented at an angle θ with respect to this line. The last processing stage consists of arithmetic averaging of responses across N such subunits, followed by a threshold. Since the individual subunit outputs are guaranteed to be non-negative, this threshold must be set at a value strictly greater than zero to have any effect.

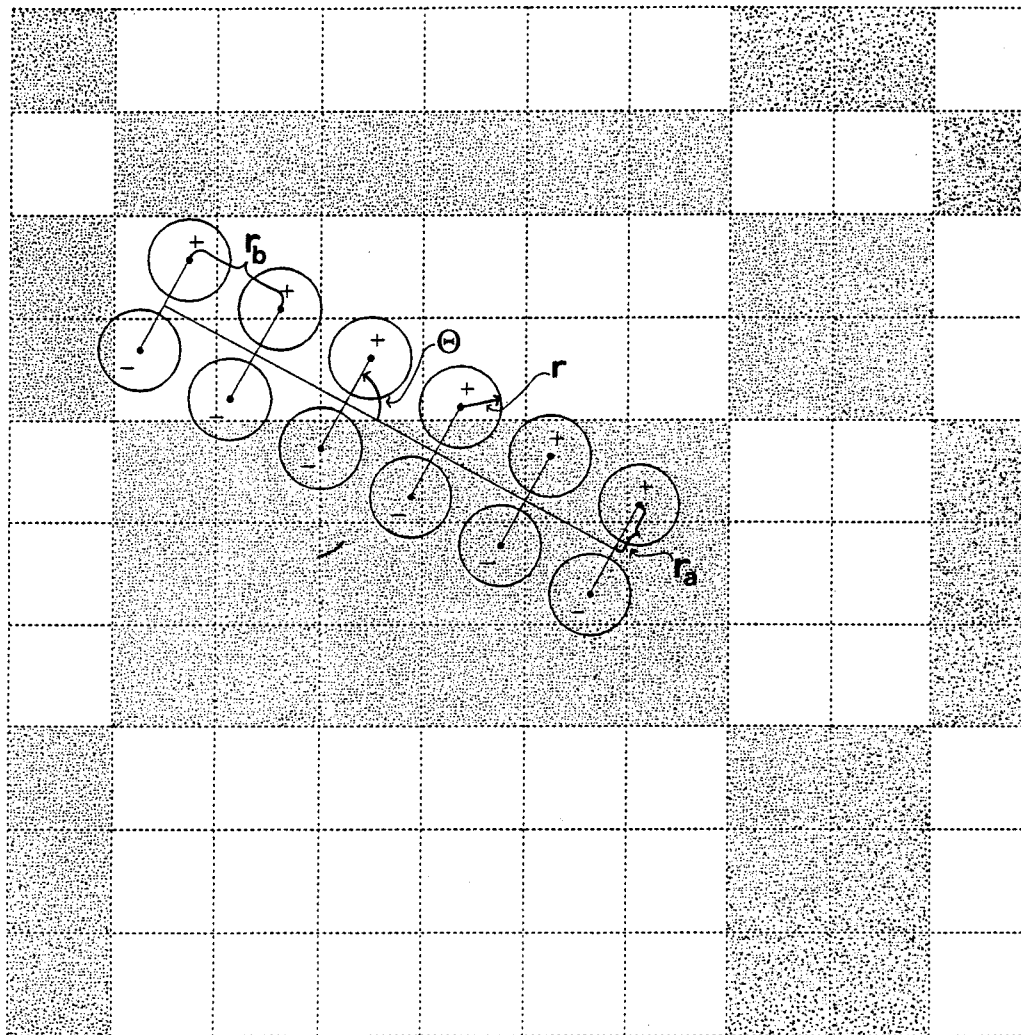


Fig. 7. A diagram of the spatial organization of the proposed model, superimposed on an example of the standard texture. Responses from circularly-symmetric linearly-summing regions are combined in an antagonistic fashion. The net response of each pair of regions is half-wave rectified. The rectified signal is averaged across pairs of regions and subjected to a thresholding operation. r is the radius of a linear-summing region, $2r_a$ is the separation of two summing regions within an antagonistic pair, and r_b is the separation of summing regions along the longitudinal axis. The number of pairs of regions is N , the threshold per pair of regions is h , and the angle between the region pairs and the longitudinal axis is θ . In this example, $r = 0.4$, $r_a = 0.5$, $r_b = 1.0$, $N = 6$ and $\theta = 90$ deg.

We denote the threshold for the second nonlinear stage by h . Since subunit responses are averaged prior to this thresholding operation, h may be considered as the threshold per subunit. (We choose to average, rather than add, subunit responses so that the influence on model responses of changing the number of subunits N is not confounded with an effective change of the threshold.)

We now examine the dependence of the model response on its defining parameters. To simulate model responses, the computational element described above was placed in at least 10 random orientations and positions on each

of at least 100 examples of the textures under consideration. Typically, these 1000 placements provided Monte Carlo estimates which were stable to within 5%. If greater variability was encountered, or if greater accuracy was required for comparisons, the Monte Carlo estimate was extended to include further placements. The average response over all such placements will be presented here.

Since the most important part of the model is the second nonlinear stage, we examine the effect of h first (Fig. 8A). The spatial parameters of the model are taken to be $N = 6$, $r = 0.5$, $r_a = 0.5$, $r_b = 0.5$, $\theta = 90$ deg. In comparing

responses to the random texture and the standard (even) texture, there are two regimes of behavior. For low values (0.0–0.2) of the threshold per subunit h , the response to the standard texture is *less* than the response to the random texture, and the difference is small (0.97:1). As the threshold parameter h increases beyond 0.4, responses to the standard texture

exceed responses to the random texture by progressively larger ratios. However, the overall response size becomes progressively smaller. Thresholds in the range 0.4–0.8 provide standard:random response ratios of 1.1–1.6, with response attenuations (compared to $h = 0$) of 73–95%. The low- h regime is unattractive since there is only a slight difference between responses to standard and random textures. We therefore choose $h = 0.6$ for further exploration.

Figure 8B shows model responses as a function of the number of subunits N . The threshold parameter h is kept fixed at 0.6 and the spatial parameters of the model unchanged from Fig. 8A. When N is small (1 or 2), responses to standard and random textures are nearly equal (standard:random response ratio of 1.05:1 or less). As N increases, this response ratio grows while response size diminishes. For $N = 16$, the response ratio is 1.7:1, but overall response size has been attenuated by approx. 90%.

The tradeoff of selectivity (large standard:random response ratio) for response size is a general feature of this model, and is not restricted to these specific values of N and h . Intuitively, this may be understood as follows. One salient feature of the standard texture is that it contains a larger-than-chance number of extended contours. Because of this, when one subunit is well-stimulated, there is a greater chance that another one along the same line is also well-stimulated. The second nonlinearity (the threshold controlled by the parameter h) results in a cooperative interaction between the subunits: there is a net model response only when several of the subunits are simultaneously active. As h increases, a greater proportion of subunits must be simultaneously active, or the levels of activity in the subunits must be higher. As N increases, the simultaneous activity of subunits must extend over a greater range (since the threshold level is set at Nh).

The $N = 6, h = 0.6$ pair leads to a reasonable selectivity for the standard texture of 1.3:1. The overall response size is approx. 30% of the response at $N = 1$, and approx. 10% of the response at $h = 0$. The reduction in responsiveness due to the second nonlinearity appears comparable to the penalty that cortical units appear to pay for selectivity in other circumstances. For example, the macaque cortical neurons have orientation tuning bandwidths ranging from approx. 5 to 60 deg, and typically have only a very small response to all but 10

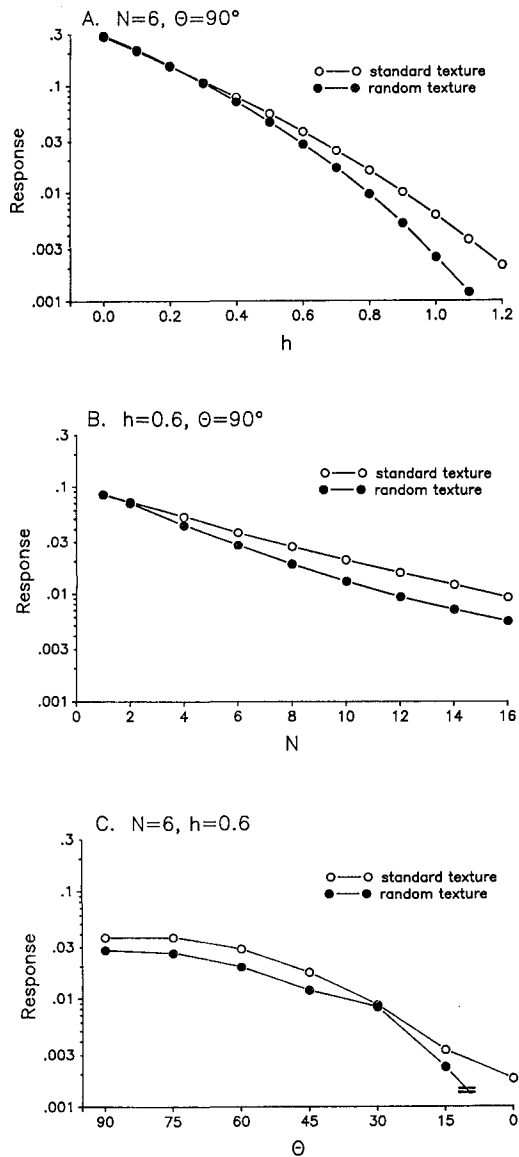


Fig. 8. Part A: simulation of model responses to the standard (even) and the random textures, as a function of the threshold parameter h . $N = 6, r = 0.5, r_a = 0.5, r_b = 0.5, \theta = 90$ deg. Part B: simulation of model responses to the standard and the random textures, as a function of the number of subunits N . $h = 0.6, r = 0.5, r_a = 0.5, r_b = 0.5, \theta = 90$ deg. Part C: simulation of model responses to the standard and the random textures, as a function of the orientation θ . $N = 6, h = 0.6, r = 0.5, r_a = 0.5, r_b = 0.5$. (○) response to standard texture; (●) response to random texture.

to 50% of possible orientations (De Valois, Yund & Helper, 1982).

Part C of Fig. 8 examines the effect of the angle θ between the edge detector axis and the line connecting subunit centers. For values of θ from 45 to 90 deg, there is relatively little effect on the standard:random response ratio. The miniscule response size for values of 15 deg or less make determinations of the response ratio meaningless in this range. Thus, orientations θ other than 90 deg provide little if any increase in the standard:random response ratio, and orientations less than 30 deg produce an unreasonably large reduction in response size.

We now examine the response of this model with $N = 6$, $h = 0.6$, $\theta = 90$ deg to the several classes of isodipole textures we have studied experimentally. As shown in the final column of Table 6, responses to the Group I textures are in the range 0.0292–0.0348, while responses to the Group III textures are in the range 0.0242–0.0279, and the response to the Group II textures (independent of polarity) is 0.0289. The response to the random texture (not listed in Table 6) is 0.0283. All Group I textures yield a larger model response than the random texture, and all Group III textures yield a smaller response than the random texture. Within Group I, the response to the cross texture is the smallest, and the response to the standard texture is the largest. The Group II response is larger than the response to a random texture, but smaller than the Group I responses. Thus, this receptive-field model indeed separates the readily-discriminable textures from those that are not.

Essential features and limitations of the model

In the data presented, we chose only a fixed set of values for the summing radius r and the spatial parameters r_a and r_b of the model. The value $r = 0.5$ for the summing radius was chosen so that the summing diameter matched the check size, and the spacing parameters were chosen so that typical summing circles overlapped only partially, and typically fell in neighboring checks. For values of r , r_a and r_b within a factor of two or so of these choices, we found similar separation of responses to textures in the different groups for appropriate choices of N and h , although model responses were generally smaller. For example, separating the subunits by $r_b = 1.0$ checks rather than $r_b = 0.5$ checks yielded responses to the Group I textures in

the range 0.0096–0.0150, responses to the Group III textures in the range 0.0074–0.0086, a response to the Group II textures of 0.0107, and a response to the random texture of 0.0102 with $N = 6$, $h = 0.6$. Within groups, the rank-order of response size depended in a complex and sensitive manner on these parameters, except that usually, the standard texture produced the largest Group I response, the cross texture produced the smallest Group I response, and the wye texture produced the smallest Group III response. In view of the variability of rank-ordering across subjects, it is difficult to use this information to refine the spatial structure of the model further.

The critically important feature of the model is the presence of *two* nonlinearities. The first nonlinearity is required so that the model's response to edges will not cancel each other if the polarity of the edges are opposite. The specific nonlinear function is not important. We have chosen half-wave rectification, but full-wave rectification, squaring, or any nonlinearity with even-order components will suffice to prevent cancellation of responses to opposite-polarity edges.

The second nonlinearity is required to generate a cooperative interaction between its local inputs. Removing the effect of the second nonlinearity (either by setting $h = 0$ or $N = 1$) eliminates the differential response to the structured textures. The second nonlinearity was chosen to be a threshold set at a level modestly higher than its typical input. This effectively provided a cooperative interaction between its inputs. It is the cooperativity, rather than the threshold itself, which is important for model behavior; other accelerating nonlinearities would serve as well.

The important spatial feature of the linear stage of the model is that it is high-pass, or relatively sensitive to edges. We chose a difference of offset Gaussians as the initial linear stage of the model to provide sensitivity to edges in a simple fashion, and to be able to study the importance of the orientation of this filtering stage to the model axis (θ). Other than this difference, similar results would have been obtained for other filters which are high-pass, such as a center/surround difference of Gaussians. The linear subunits of the model were arrayed along a line because this arrangement provided the best fit to VEP responses to standard textures with sporadic and propagated decorrelation (Victor & Conte, 1989).

Aspects of the model which are incomplete

The model as it stands might provide an adequate explanation of responses at a given contrast level, but cannot account for responses over a range of contrasts. This is because the effects of the second nonlinearity become progressively weaker as contrast increases. Indeed, since the local nonlinearity (summation in circular regions followed by half-wave rectification) has a response which is proportional to stimulus contrast, scaling the contrast up by a factor c is equivalent to scaling the threshold down by a factor c . In descriptive terms, the selectivity which was conferred by the thresholding operation can be defeated simply by presenting a stimulus with sufficiently high contrast. Contrary to this model prediction, pattern-specific responses to isodipole textures do not decrease (either in absolute or relative size) as contrast increases (Victor & Conte, 1987). This difficulty is not unique to this model, but would be shared by any model which gains selectivity through a cooperative interaction. A reasonable extension of the model which solves this problem is that the threshold is adjusted according to the global contrast level (e.g. through lateral inhibitory interactions that are not specific to stimulus features).

We have not attempted to model temporal aspects of processing. It does not suffice to regard the model's output as a scalar, to which appropriate dynamics which generate the VEP waveform may be attached. Model responses to Group I and Group II textures are *larger* than responses to the random texture, while responses to Group III textures are *smaller* than responses to the random texture. If the VEP simply represented the output of the spatial model shaped by a temporal filter, one would predict that Group III responses would be opposite in phase to Group I responses, rather than much smaller but of a similar phase, as we have found. A more complex temporal model (perhaps including dynamical components at either or both of the postulated nonlinear stages) is therefore required.

We have also not accounted for the subtle but definite difference in the dynamics of responses to different textures. It is unclear whether this difference can be accounted for simply by adding the appropriate dynamics to the proposed model. An alternative and more attractive explanation rests on the statistical properties that distinguish Group II textures from Group I

textures. A Group II texture contains figures of a particular contrast polarity (bright or dark, but not both) on a random-appearing background. Interchange of a Group II texture with the random texture may therefore yield a VEP because of asymmetries between "on" and "off" pathways (Zemon, Gordon & Welch, 1988). However, Group I and random textures cannot probe on/off asymmetries: inverting the contrast of a Group I or a random texture merely produces another example of the same texture. Thus, the difference between Group I and Group II responses may well be due to the existence of neural mechanisms driven by Group II textures that are not driven by the random or Group I textures. This is suggested by two pieces of evidence. (i) The analysis of the invariant $I(G, G')$ demonstrated that the need for a second principal component was primarily due to a comparison of Group I with Group II responses. (ii) Model responses for Group II textures are smaller than those for Group I textures, but VEP and psychophysical responses are similar in size.

Relationship to cellular studies of cortical information processing

Despite this model's incompleteness in several respects, it raises basic questions concerning the computations of early vision. Models for simple cell receptive fields are typically linear, or, linear with a single stage of rectification at the stage of spike generation (Movshon et al., 1978a; Camarda et al., 1985; Spitzer & Hochstein, 1985a; Jones & Palmer 1987). The complex cell receptive field has been modelled as multiple nonlinear subunits arrayed on a line, whose outputs are summed linearly (Spitzer & Hochstein, 1985b). We find that a second nonlinearity and a stage of spatial pooling is crucial to account for the difference in response to structured and random textures. As seen in Fig. 8A, removal of the second nonlinearity (setting $h = 0$) eliminates the difference in response to the structured and random textures. As seen in Fig. 8B, removal of the final stage of spatial pooling (setting $N = 1$ or $N = 2$) also eliminates this difference. Thus, if the single-nonlinearity concept of simple and complex receptive field structure is accurate, we can infer that the computations we have been probing take place at a later stage, perhaps by cooperative combination of simple or complex cell responses.

The problem with this line of reasoning is that the presence of the second nonlinearity in the current model only makes a qualitative difference for intrinsically two-dimensional stimuli, such as the textures we have used. The effects of the second nonlinearity would not be very evident if the test stimuli used were one-dimensional stimuli, such as gratings and bars, even if presented at many orientations. This can be seen without detailed modeling, as follows. Oriented along the preferred axis of the cell, a bar or grating would excite all the subunits in synchrony, and the two nonlinearities would collapse into a single threshold. Oriented along a nonpreferred direction, such a stimulus would evoke little response from the first stage of the model, and again the effects of the nonlinearity would not be evident.

Thus, it is possible that the two-nonlinearity network needed to account for our VEP and psychophysical results is present at the level of the simple and complex cell, but that most studies based on traditional stimuli do not reveal its presence. The possibility of a cooperative interaction among subunits of a complex cell is raised by the work of Movshon et al. (1978b) based on responses to asynchronously-presented pairs of line segments; our model requires a cooperative interaction between synchronously-presented elements as well. If indeed the second nonlinearity resides early in striate circuitry, then the connections between subunits may be related to the long-range excitatory and orientation-specific connections in striate cortex (Ts'o, Gilbert & Wiesel, 1986).

Relation to models of texture segmentation

Our psychophysical task is discrimination of sequentially-presented homogeneous textures, rather than one of texture segregation. But since early vision mechanisms that are involved in texture discrimination are likely related to those involved in texture segregation, we consider the relationship of our model to models proposed for texture segregation.

Beck, Sutter and Ivry (1987) interpreted the cancellation of opposing changes in size and contrast as evidence that texture segmentation is based on the output of spatial frequency channels. The present work, as well as the studies of Julesz and coworkers (Caelli & Julesz, 1978, Caelli, Julesz & Gilbert, 1978; Julesz, 1981), reveals that local features support texture discrimination even under circumstances in which the output of spatial frequency channels

are statistically identical. Thus, models based on spatial frequency channels cannot account for our results.

Bergen and Adelson (1988) discuss a class of models for texture segmentation based on local energy measures. In these models, local energy is computed by rectification of the output of localized spatial filters. Detection of texture differences is based on the regionally summed outputs of these local filters. Chubb and Landy (1990) elaborate on this model by considering a single nonlinear stage of more arbitrary characteristics. But since these models lack a second nonlinear stage, they are unlikely to explain our findings without further elaboration.

Such elaboration may be found in the models of Caelli (1985) and Voorhees and Poggio (1988), who incorporate local energy detectors in a more complex framework which includes long-range nonlinearities. In these models, the primary purpose of the additional nonlinearities is to account for processes such as filling-in and boundary generation, which are considered to follow initial feature extraction. Though these spatial nonlinearities are different from the second nonlinear stage that we propose, it is interesting to note that they (or similar processes) appear necessary to account for aspects of the initial analysis of form as well.

Acknowledgements—This work was supported in part by grants EY1428 and EY7977.

REFERENCES

- Beck, J., Sutter, A. & Ivry, R. (1987). Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics and Image Processing*, 37, 299–325.
- Bergen, J. R. & Adelson, E. H. (1988). Early vision and texture perception. *Nature, London*, 333, 363–364.
- Caelli, T. (1985). Three processing characteristics of visual texture segmentation. *Spatial Vision*, 1, 19–30.
- Caelli, T. M. & Julesz, B. (1978). On perceptual analyzers underlying visual texture discrimination: Part I. *Biological Cybernetics*, 28, 167–175.
- Caelli, T. M., Julesz, B. & Gilbert, E. N. (1978). On perceptual analyzers underlying visual texture discrimination: Part II. *Biological Cybernetics*, 29, 201–214.
- Camarda, R. M., Peterhans, E. & Bishop, P. O. (1985). Spatial organization of subregions in receptive fields of simple cells in cat striate cortex as revealed by stationary flashing bars and moving edges. *Experimental Brain Research* 60, 136–150.
- Cattell, R. B. (1978). *The scientific use of factor analysis*. New York: Plenum Press.
- Chubb, C. & Landy, M. S. (1990). Orthogonal distribution analysis: A new approach to the study of texture perception. In Landy, M. S. & Movshon, J. A. (Eds),

- Computational models of visual processing*. Cambridge, Mass.: MIT Press.
- Desimone, R., Schein, S. J., Moran J. & Ungerleider, L. G. (1985). Contour, color and shape analysis beyond the striate cortex. *Vision Research*, 25, 441–452.
- De Valois, R. L., Albrecht, D. G. & Thorell, L. G. (1978). Cortical cells: bar and edge detectors, or spatial frequency filters? In Cool, S. J. & Smith, E. L. (Eds) *Frontiers in visual science*. (pp. 544–556). Berlin: Springer.
- De Valois, R. L., Yund, E. W. & Helper, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, 22, 531–544.
- Diaconis, P. & Freedman, D. (1981). On the statistics of vision: the Julesz conjecture. *Journal of Mathematical Psychology*, 24, 112–138.
- 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.
- Farnsworth, D. (1957). The Farnsworth–Munsell 100-Hue test for the examination of color discrimination: Manual. Baltimore, MD: Munsell Color Company.
- Gilbert, E. N. (1980). Random colorings of a lattice on squares in the plane. *SIAM Journal of Algebra and Discrete Methods*, 1, 152–159.
- 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.
- Hochstein, S. & Shapley, R. M. (1976). Linear and non-linear spatial subunits in Y cat retinal ganglion cells. *Journal of Physiology, London*, 262, 265–284.
- Hubel, D. H. & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology, London*, 195, 215–243.
- Jones, J. P. & Palmer, L. A. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58, 1187–1211.
- Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature, London*, 290, 91–97.
- Julesz, B., Gilbert, E. N. & Victor, J. D. (1978). Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics*, 31, 137–140.
- Milkman, N., Schick, G., Rossetto, M., Ratliff, F., Shapley, R. & Victor, J. D. (1980). A two-dimensional computer-controlled visual stimulator. *Behavioral Research Methods and Instrumentation*, 12, 283–292.
- Morrone, C., Burr, D. C. & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells—I. Neurophysiological evidence. *Proceedings of the Royal Society, London B*, 216, 335–354.
- Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. (1978a). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology, London*, 283, 53–77.
- Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. (1978b). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology, London*, 283, 79–99.
- Regan, D. (1989). *Human brain electrophysiology*. New York: Elsevier.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Systems Technical Journal*, 27, 379–423.
- Spitzer, H. & Hochstein, S. (1985a). Simple- and complex-cell response dependences on stimulation parameters. *Journal of Neurophysiology*, 53, 1244–1265.
- Spitzer, H. & Hochstein, S. (1985b). A complex-cell receptive-field model. *Journal of Neurophysiology*, 53, 1266–1286.
- Szulborski, R. G. & Palmer, L. A. (1990). The two-dimensional spatial structure of nonlinear subunits in the receptive fields of complex cells. *Vision Research*, 30, 249–254.
- Ts'o, D. Y., Gilbert, C. D. & Wiesel, T. N. (1986). Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *Journal of Neuroscience*, 6, 1160–1170.
- Victor, J. D. (1985). Complex visual textures as a tool for studying the VEP. *Vision Research*, 25, 1811–1827.
- Victor, J. D. (1988). Models for preattentive texture discrimination: Fourier analysis and local feature processing in a unified framework. *Spatial Vision*, 3, 263–280.
- Victor, J. D. & Conte, M. M. (1987). Visual evoked potentials elicited by simple and complex textures: Distinct components with similar scalp topographies. In Barber, C. & Blum, T. (Eds), *Evoked potentials III: The Third International Evoked Potentials Symposium* (pp. 183–189). Boston: Butterworths.
- Victor, J. D. & Conte, M. M. (1989). Cortical interactions in texture processing: scale and dynamics. *Visual Neuroscience*, 2, 297–313.
- Victor, J. D. & Conte, M. M. (1990). Motion mechanisms have only limited access to form information. *Vision Research*, 30, 289–301.
- Victor, J. D. & Zemon, V. (1985). The human visual evoked potential: Analysis of components due to elementary and complex aspects of form. *Vision Research*, 25, 1829–1842.
- Von der Heydt, R., Peterhans, E. & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, 224, 1260–1262.
- Voorhees, H. & Poggio, T. (1988). Computing texture boundaries from images. *Nature, London*, 333, 364–367.
- Zemon, V., Gordon J. & Welch, J. (1988). Asymmetries in ON and OFF visual pathways of humans revealed using contrast-evoked cortical potentials. *Visual Neuroscience*, 1, 145–150.

APPENDIX

Asymptotic Analysis of Number of Order-4 Correlations

For any particular glider, correlations of order 4 exist on any set S whose generating function $\hat{S}(x, y)$ is divisible by the generating function of the glider $\hat{G}(x, y)$. Here we analyze the asymptotic behavior of the number of such sets S , as a function of the distance D spanned by the set, for the gliders of Group I. We content ourselves with estimating the form of the leading term of dependence on the distance D , and omit details needed to calculate lower-order terms and numerical factors.

The simplest case to consider is that of the standard glider, which has a generating function.

$$\hat{G}_{\text{standard}}(x, y) = (1+x)(1+y). \quad (\text{A1})$$

We have seen (example following equation 8) that all polynomials of the form $\hat{S}(x, y) = (1+x^l)(1+y^l)$ have $\hat{G}_{\text{standard}}$ as a factor. We now show that these are the only four-term polynomials whose lowest-order term is 1 that have $\hat{G}_{\text{standard}}$ as a factor.

Assume that $\hat{P}(x, y)$ is a four-term polynomial with $\hat{G}_{\text{standard}}(x, y)$ as a factor. We note that:

$$\hat{G}_{\text{standard}}(x, 1) = 0 \pmod{2}. \quad (\text{A2})$$

Since $\hat{P}(x, y) = \hat{G}_{\text{standard}}(x, y) \cdot \hat{Q}(x, y) \pmod{2}$ for some quotient generating function \hat{Q} , it follows that:

$$\hat{P}(x, 1) = 0 \pmod{2} \quad (\text{A3})$$

According to the definition of the generating function (7), eqn (A3) expands to

$$\sum_{(I_s, J_s) \text{ in } P} x^{I_s} = 0 \pmod{2} \quad (\text{A4})$$

Equation (A4) is an identity between polynomials in x , with coefficients interpreted modulo 2. It can only hold if the coefficients of x^i are all *even*. This means that the number of checks (I_s, J_s) in every row is an even number. Thus, any nonempty set P whose generating function is divisible by \hat{G} must have at least two columns that contain checks. A similar analysis, starting with $\hat{G}_{\text{standard}}(1, y) = 0 \pmod{2}$, leads to the conclusion that P must have at least two rows that contain checks. It follows that if P contains only four checks, then these checks must be at the corners of

a rectangle. Assuming that the lower-left corner of this rectangle is at $(0, 0)$, then the generating function must be of the form $(x^I + 1)(y^J + 1)$.

It is now straightforward to count the number of four-element sets within a distance D that have non-zero fourth-order interactions. This is simply the number of rectangles whose dimensions are both less than D , and is proportional to D^2 .

The other Group I textures may be handled in an identical fashion. This follows from the observation that the other Group I gliders are a kind of linear transformation of the standard glider. In terms of generating functions, this corresponds to an algebraic relationship between the other Group I gliders' generating functions and $\hat{G}_{\text{standard}}$ (equations 19–21).

Because of these algebraic relationships, the divisibility arguments that applied to $\hat{G}_{\text{standard}}$ apply to the other generating functions as well, and the number of sets S whose generating functions are divisible by these generating functions also grow in proportion to D^2 . The proportionality constant depends on the shape of the glider, and the shape of the area of the texture under consideration. For typical regions, the proportionality constant is largest for $\hat{G}_{\text{standard}}$ and smallest for \hat{G}_{cross} .