

Title:

DISCRIMINATION OF ISOTRIGON TEXTURES (36 Characters)

Running Title:

Discriminating Isotrigon Textures (33 characters)

Text: 7918 words

Artwork: 17 Figures, 4 Tables

Key words: texture, isotrigon, isodipole, Allan variance

Author details:

T. Maddess Ph.D.,¹ Y. Nagai Ph.D.,²

¹Centre for Visual Sciences, RSBS, ANU, Canberra; ²Center for Information Science, Kokushikan University, Tokyo

Proprietary interest: nil

All communications to:

Dr. T. Maddess
Centre for Visual Sciences
Research School of Biological Sciences
Australian National University
Canberra ACT 0200
Australia

Tel: 612 6249 4099

Fax: 612 6249 3808

Email: ted.maddess@anu.edu.au

Abstract (100 words)

Higher order correlations can capture edge and object relationships. Isotrigon textures are a useful tool for studying our sensitivity to these correlations. We determined human discrimination performance for 18 isotrigon texture types and compared it with outputs from statistical discriminant models. Some of the models employed versions of the Allan Variance in receptive field outputs. Physiologically plausible mechanisms for such calculations are presented. Two discriminant models emulated human performance well, one based upon a global variance measure, and the other based upon a localised variance with an orientation bias. The 18 texture types were also shown to contain characteristic mini-textures.

Introduction

Beck (1987) showed that a trade-off between texture element size and contrast that affects texture segregation (Beck 1983) could be explained by spatial summation within oriented narrow-band spatial frequency channels whose output was later full-wave rectified (Sutter, Beck & Graham 1989) (by full-wave rectification we mean that the units give similar responses to contrasts that are brighter or darker than average). These so-called “complex channels”, are now well accepted (Landy & Bergen 1991; Graham, Sutter, Venkatesan & Humaran 1992; Graham & Sutter 1996; Graham & Sutter 1998; Graham & Sutter 2000). Experiments on texture defined motion also indicate the presence of a “texture grabber” based on rectified, oriented spatial filters (Chubb & Sperling 1991), where the extracted textural information may form the input to the same motion detecting mechanism as that for luminance defined motion (Turano & Pantle 1989; Werkhoven, Sperling & Chubb 1993; Werkhoven, Sperling & Chubb 1994). Nonlinear mechanisms also appear to underlie texture discrimination on the basis of orientation modulation (Kingdom, Keeble & Moulden 1995).

Julesz (1973) hypothesised that humans could not discriminate briefly presented textures whose second order correlation functions were the same, so called *isodipole* textures. While Julesz and colleagues used slightly different statistics to the standard autocorrelation function the two statistics are the same providing the compared textures have the same mean luminance (Klein & Tyler 1986). This idea seemed to hold up until a class of patterns were discovered, groups of which, although easily discriminable, have identical average third order correlation functions (Julesz, Gilbert & Victor 1978), *i.e.* correlation functions based on triplets of image points. Such *isotrigon* textures by definition also have identical means and second order correlations, that is, they are also isodipole (Gilbert 1980), and a variety of such textures have now been found (Gilbert 1980; Victor & Conte 1991). While the third order correlation function of any monochromatic image uniquely determines that image (Yellott

1993) ensemble averages of groups of isotrigon textures have third order correlation functions that are also indiscriminable from that of binary noise (Victor 1994). Hence, the average third order correlation functions of collections of isotrigon textures cannot be used to discriminate other collections of differing types of isotrigon textures. It is worth noting that ensembles of several textures need to be averaged to achieve good isotrogonicity (Gagalowicz 1981).

Victor and Conte (1991) have constructed a parsimonious neural model whose performance matches some features of human visual discrimination of these textures. The model accounts for human visual evoked responses, those responses being dependent on the higher-order statistics of the textures. That model contains neurones similar to the complex channels described above, and involves a spatial sum of an array of oriented spatial filters that have their individual responses rectified, followed by a strong threshold nonlinearity. Thus, the model mimics some of the features of complex the channels posited by others to explain other forms of texture segregation.

Purpura et al. (1994) noted that while two point correlations inform the visual system about the (phase independent) spatial frequency content of images, encoding features such as contours requires mechanisms sensitive to correlations involving three or more points. VEP studies with isotrigon stimuli have shown that the visual cortex does utilise higher order correlations (Victor 1985; Victor & Zemon 1985; Victor & Conte 1989; Victor & Conte 1991). Even some single units in macaque V1 have been shown to encode such high order statistical information, and this ability is not dependent on cortical layer (Purpura *et al.* 1994). PET (Beason-Held, Purpura, Van Meter, Azari, Mangot, Optican, Mentis, Alexander, Grady, Horwitz, Rapoport & Schapiro 1998) and fMRI (Beason-Held, Purpura, Krasuski, Maisog, Daly, Mangot, Desmond, Optican, Schapiro & VanMeter 1998; Beason-Held, Purpura, Krasuski, Desmond, Mangot, Daly, Optican, Rapoport & VanMeter 2000) studies have revealed brain regions specialising in discriminating random from isotrigon patterns.

The isotrigon textures used to date each have two related forms, the so-called Even and Odd variants. The studies mentioned above have used one isotrigon texture class, the Even and Odd patterns generated by the 2x2 Box glider (Fig. 1, column 1) and some variants. One study (Victor & Conte 1991) has examined discrimination of Random binary textures from modified versions of the 9 Even textures used here. Given the interesting possibilities for using isotrigon textures in studying vision we sought to quantify discrimination between 27 combinations of Even *vs.* Random, Odd *vs.* Random, and Even *vs.* Odd textures. These discriminations were quantified for four octaves of texture sizes. We also examined the effect of texture contrast and scale, random number seed, and examined long-term learning effects. Since we were studying discrimination we also sought to compare human performance with that produced by formal statistical discriminant models that shared some features of common texture segregation mechanisms.

{Figure 1 about here please}

Since a quantity known as the Allan Variance has computational similarity to physiological mechanisms for calculating even order nonlinear interactions we explored the use of four versions of Allan Variance as inputs to the discriminant models. These variance measures were based on outputs of localised receptive field models. Linear and quadratic discriminant models were considered. Since the variance models supplied second order statistical information the resultant linear and quadratic discriminant models were formally comparing second and fourth order spatial correlations.

We also examined the local properties of the textures by characterising the number of types of small 3x3 and 4x4 pixel sub-domains of the textures, which we refer to as mini-textures. We related the characteristic number and form of mini-textures found within each class of the 18 texture types we investigated.

Methods

Stimuli

Achromatic (colour temperature 6500 °K) texture patterns were displayed at mean luminance of 45 cd m^{-2} on a Barco CCID 7551 monitor at a pixel resolution of 512 x 420 square pixels, and a refresh rate of 101 Hz. The ambient illumination was that provided by the monitor in an otherwise darkened room. Monitor linearity was confirmed by nonlinear systems identification methods (Maddess, James, Goldberg, Wine & Dobinson 2000). Subjects viewed the patterns binocularly from a distance of 60 cm with the aid of a chin rest.

The texture patterns used here include random binary textures and the 9 classes described by Victor and Conte (1991). For the *random* patterns the probability of a given check being dark or bright was 0.5. The non-random textures were created by a recursion process (Victor & Conte 1991). To begin the process a matrix is set randomly dark (-1) or bright (1) with probability 0.5. Then checks are then recoloured in by a recursion rule operating *via* a moving 3x3 pixel domain. The recursion rule only operates on a subset of the 9 pixels, the shape of these active pixels being referred to as a *glider*. Nine different gliders define the 9 textural types used here (Fig. 1). Each glider can generate two related types of texture known as the *Even* and *Odd* versions. During the recursion the product of the pixels defining the glider shape is held to be 1 for Even textures, -1 for Odd. For example, to produce Even textures from 4 pixel gliders the rule can be achieved by setting the fourth pixel to equal the product of the other 3, and thus the subsequent product with the 4th pixel is always 1. By inverting the sign of the product of the first 3 glider elements, and setting the fourth pixel equal to that, the product of the 4 glider pixels will thus always be -1. In either case the product of the first three pixels determines the value of the fourth. The statistics of the generated pattern does not depend on the initial pixel values or in which direction the recursion is operated (Gilbert 1980). With the exception of the Triangle texture, ensemble

averages of all these texture types have the same third order correlation functions as that for binary noise patterns, and are thus said to be isotrigon textures (Victor 1994). By definition this means the patterns also have equal second order correlation functions and are thus isodipole textures. Blurring of the patterns can foil their isotrigon property but does not affect their isodipole property since blurring will have the same affect on the 2-D power spectra, which are the Fourier transforms of the 2-D correlation functions (Victor 1994). The Triangle textures are only isodipole even without blurring and so they provide an interesting contrast.

Four sizes of textures were employed: 4x4, 8x8, 16x16 and 32x32 pixels. Typically the pixel blocks were 58' square measured at the CRT centre. Some trials employing textures made with 4 times smaller blocks (14.5' square) were also conducted. The large block pixels used here minimised the possible effects of blurring. All textures were presented at the CRT centre, the remaining 32 x 24 deg (h x v) face of the CRT was a blank field at the mean luminance. Within each presentation the contrast of the gratings was increased from 0 to the test contrast and then back down to 0. To minimise the effect of onset and offset transients the leading and trailing 5 frames of each presentation followed a Blackman function: $\text{Blackman}(t) = 0.42 - 0.5 \cdot \cos(2\pi(t-t_0)/\tau) + 0.08 \cdot \cos(4\pi(t-t_0)/\tau)$ (Fig. 2). The stimulus was thus on at full contrast for 204 ms and the total stimulus duration was 297ms (Figure 2). In most experiments textures were presented at contrast 1 but in some trials contrasts of 0.1 or 0.2 were used.

{Figure 2 about here please}

For a particular block of trials subjects were shown either Even or Random textures, Odd or Random textures, or Even or Odd textures, where the non-random textures were drawn using one of the 9 gliders and the appropriate Odd or Even rule. These three types of discriminations are henceforth referred to as ER, OR and EO discriminations. Before each block of trials subjects were given printed black and white examples of the textures having

64x64 pixels (15x15 cm) to inspect for up to 3 minutes. Within a given block of trials 12 textures samples of different sizes were presented: first at 32x32 pixels, then 12 textures at 16x16 pixels and so on. Subjects indicated by a mouse button press whether a given texture was Even or Random, Odd or Random, or Even or Odd, depending on the type of trial being conducted. Incorrect choices were indicated by a tone. Blocks of trials for each glider texture class were completed in the order ER, OR, EO. Blocks of trials for the textures generated by the 9 different gliders were completed at random providing 27 (x12) discriminations each for up to 4 texture sizes. Two subjects (TM, YN) completed 5 repeats of the entire process, partly to study long term learning, and partly to test a few subjects with textures generated from several different random number seeds.

Statistics

We employed both linear and quadratic discriminant analyses (Johnson & Wichern 1992) to attempt to mimic the human discrimination process within a statistical formalism. In the linear discriminant case the covariances \mathbf{S} are assumed to be equal and are pooled, $\mathbf{S}_{\text{pooled}} = ((m-1)\mathbf{S}_1 + (n-1)\mathbf{S}_2) / (m+n-2)$ to obtain a better estimate. If we have two multivariate normal populations π_1 and π_2 the condition for assigning an observation \mathbf{x}_0 to π_1 , for the case of equal prior odds of encountering π_1 or π_2 and equal costs for mis-classification as π_1 or π_2 , is:

$$(1) \quad (\mathbf{u}_1 - \mathbf{u}_2)' \mathbf{S}^{-1} \mathbf{x}_0 + \frac{1}{2} \mathbf{u}_2' \mathbf{S}^{-1} \mathbf{u}_2 - \frac{1}{2} \mathbf{u}_1' \mathbf{S}^{-1} \mathbf{u}_1 > 0$$

Where the \mathbf{u}_n are the means for π_1 and π_2 , \mathbf{S} is the pooled covariance, and \mathbf{x}_0 is a sample (vector) to be classified as either being from π_1 and π_2 . This condition can be written in the form

$$(2) \quad \mathbf{b} \mathbf{x}_0 + \mathbf{c} > 0$$

where \mathbf{c} is a scalar, and \mathbf{b} is a row vector of coefficients. This is sometimes referred to as Fischer's *linear* discriminant or classification rule.

If the covariance matrices are not equal, i.e. $\mathbf{S}_1 \neq \mathbf{S}_2$, then the resultant rule is to classify an observation \mathbf{x}_0 as being from population 1 (π_1) if the following quadratic condition holds:

$$(3) \quad \mathbf{x}_0' \mathbf{A} \mathbf{x}_0 + \mathbf{b}' \mathbf{x}_0 + \mathbf{c} > 0$$

where

$$\mathbf{A} = \frac{1}{2} (\mathbf{S}_2^{-1} - \mathbf{S}_1^{-1})$$

$$\mathbf{b} = -\frac{1}{2} (\mathbf{u}_2' \mathbf{S}_2^{-1} - \mathbf{u}_1' \mathbf{S}_1^{-1})$$

$$\mathbf{c} = \frac{1}{2} (\mathbf{u}_2' \mathbf{S}_2^{-1} \mathbf{u}_2 - \mathbf{u}_1' \mathbf{S}_1^{-1} \mathbf{u}_1) + \frac{1}{2} \log(\det(\mathbf{S}_2)/\det(\mathbf{S}_1))$$

Geometrically, the linear discriminant function defines a decision boundary, or *separatrix*, that is a (possibly diagonal) straight line across the plane in the case of a bivariate observation, and in general is a hyperplane in the multivariate case. The quadratic discriminant rule defines a decision boundary that is (possibly rotated) parabola in the bivariate case, and is a paraboloidal surface in the general multivariate case (for examples see (Maddess, Goldberg, Wine, Dobinson, Welsh & James 1999; Maddess *et al.* 2000)).

Subjects and general

The authors and three naive subjects were used. All subjects were male and ranged in age between 22 and 49 years and had normal or corrected-to-normal vision. This study was conducted under protocol M881 of the Human Ethics Experimentation Committee of the Australian National University. All calculations were performed using Matlab Release 11 (The Mathworks, Natick MA), and the random number seeds were set by using the *state* variable of its “rand” function.

{Figure 3 about here please}

Results

Mini-textures

As a starting point it is reasonable to assume that the visual system may initially perform a local analysis of the textures. This leads to consideration of small two-dimensional domains of the parent textures that we refer to here as mini-textures. For an $N \times M$ pixel binary random mini-texture the possible number of mini-textures is 2^{NM} , hence for a 3×3 domain we obtain $2^9 = 512$, and for 4×4 pixels $2^{16} = 65536$. It can be shown (Appendix) that for textures like those used here, where the number of active pixels of the glider is less than the total number of pixels in the whole glider domain (3×3 in the present case), that the total number of mini-textures generated by the recursion process will be less than 2^{nm} .

{Tables 1 and 2 about here please}

As an empirical check of this idea we took 16 Odd and 16 Even 100 pixel² textures for each of our 9 texture/glider types and examined the unique mini-textures in each (*e.g.* Fig. 2a,b). Each of the 16 examples of the resulting 18 texture types was constructed from a different random number seed. The $\{-1,1\}$ textures were converted to $\{0,1\}$, the columns were concatenated to a single vector, and the resulting binary number was used as an index identifying each unique mini-texture type. The indices were used to create histograms of the number of mini-textures in a given parent texture (*e.g.* Fig. 3c). Notice that for a 3×3 mini-texture there are thus $2^9 = 512$ indices, *i.e.* one histogram bin for every possible 3×3 mini-texture. The resultant counts were used to form frequency histograms (*e.g.* Fig. 3d) of the number of occurrences of each mini-texture in much larger parent textures. The process was repeated for 4×4 mini-textures. In this case the 288 (18 types \times 16 examples) textures were used, each 416 pixel² to insure a sufficient number of each mini-texture was obtained for each of the 65536 possibly valid histogram bins. While the analysis in the Appendix predicts the number of unique mini-textures in each texture type (*cf.* Tables 1 and A1) it says nothing

about the mini-textures that are shared between the various texture types. The above empirical analysis permitted these relationships to be quantified (Table 2, and Fig. 4).

{Figure 4 about here please}

As illustrated by Figure 3d the histogram of mini-texture frequencies were quite flat but nevertheless showed a structure having mirror symmetry in the frequencies of appearance of particular textures. The same symmetry was present in the variance of the frequencies of appearance of particular mini-textures. Frequency histograms for both 3x3 and 4x4 mini-textures from all 18 texture types exhibited this mirrored property although the mirrored patterns differed for different textures. As shown in Table 3 the frequency of mini-textures was similar for Even and Odd texture pairs, but the variance in the frequencies was the same or larger for Even mini-textures. So, for example, for the 3x3 mini-textures taken from 100 pixel² Even Zigzag textures each possible mini-texture occurs 75.0 ± 27.2 SD times, and 75.0 ± 18.0 SD times for Odd Zigzag patterns.

{Table 3 and Figure 5 about here please}

Notice that all the textures should in principle be perfectly discriminable from random textures given that the *alphabet* of 3x3 mini-textures defining any of the 18 texture types used here will never have more than half the *letters* (mini-textures) found in random binary textures. In practice human performance is less than perfect (see below). A possible source of poor performance is spatial summation by receptive fields. A trivial example would be a hypothetical receptive field that summed all the pixels within a 3x3 domain with equal weights, clearly this would produce the same output for every mini-texture having 1 bright and 8 dark pixels and so on. Thus, summation could greatly reduce the number of observable responses compared to the number of mini-textures, even without consideration of noise. Spatial summation may in effect provide humans with a reduced alphabet of mini-textures with which to discriminate binary textures.

{Figure 6 about here please}

Psychophysics

Since we were interested in the local properties of the textures under study for each texture type subjects made ER, OR and EO discriminations for 3 or 4 texture sizes (Methods) within each block of trials. In the first block of trials textures that 32x32 pixels were presented, and in subsequent trials texture size was reduced in octave steps to as small as 4x4 pixels. Figure 5 shows the average outcomes for 5 subjects. Two different random number seeds were used for these data, two subjects viewed textures produced from one seed and the other three subjects viewed textures from a second seed. Discrimination fell generally below the 75% level when texture size was reduced to 8x8 pixels (Fig. 5f), although textures 1 (Box) and 2 (Triangle) continued at near saturated performance levels. The authors completed 5 repeats for all textures each viewing textures made using 5 different random number seeds (Fig. 6). The resulting psychometric functions are very similar across subjects and seeds (*cf.* Figs. 5 & 6). The authors also repeated all the discriminations at two lower contrasts (0.1 and 0.2) and a 4 times smaller check size. Averages cross both subjects and for different cases are shown in Figure 7. The averages across discriminations (Figs 5,6,7 b,d,f) are very similar all having a dip at textures 3 (Cross), 7 (Wye) and 9 (El) (see also Fig. 16).

{Figure 7 about here please}

The authors also completed two repeats of all discriminations (24 repeats per texture and size, for ER, EO, and EO) for 4x4 pixels (Fig 8). These data were collected in blocks of trials containing the larger texture sizes obtained during the first 2 repeats of the experiments of Figure 6. Even for these quite experienced observers performance fell to 65% or lower for all but textures 1 (Box) and 2 (Triangle). A simple model is also shown together with the average discrimination data for the 4x4 pixel case (Fig. 8b). Given that Random binary 3x3 mini-textures can have 512 possible states, R, and the deterministic textures (see Appendix for

definition) we have used are constrained to a lesser number of mini-textures, N (Appendix, Table 1), then a simple idea would be that discrimination, D , might be $D = 1 - (R - N)/R$. D is plotted as a solid line in Figure 8b. While D has some features of the psychometric function there is no *a priori* reason to assume that the frequency of 3x3 mini-textures is especially relevant to discrimination of 4x4 textures.

{Figure 8 about here please}

Since the authors completed 5 repeats of the all discriminations over the space of several months it is reasonable to look for learning effects. Linear regression across repeats on data for each texture and size showed significant learning in some cases. Figure 9 summarises the regression data where the 30 data sets per subject (two authors, 5 repeats, ER, OR, EO) have been fitted together. Starred symbols were significant at $p < 0.05$. The significant learning effects were observed for textures 3 (Cross) and 9 (EI) at all sizes, and for the larger textures generally. The curves for smaller textures are similar in shape to those for the larger textures suggesting that the lower slopes may contribute to the lack of significance.

{Figure 9 about here please}

Modelling

Since the task we were examining was the ability of humans to make various textural discriminations we thought it might be useful to explore formal statistical discriminant models as a way of obtaining insights into, or quantifying constraints upon, human performance. We employed Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) (see Methods). For two populations, π_1 and π_2 , the object of discriminant analysis is to assign an observation \mathbf{x}_0 to one population or the other (Methods). In our case π_1 and π_2 represent some measures obtained from populations of textural types in an ER, EO or EO comparison. Clearly several measures on each example could be made, for example some measure for each of the 3x3 mini-textures contained in a 12x12 texture sample. Thus, \mathbf{x}_0 is in

general a vector $\mathbf{x}_0 = \{ x_1, x_2, x_3, \dots, x_n \}$. In the case of two dimensional data, $\mathbf{x}_0 = \{ x_1, x_2 \}$, where the observations and populations span a plane, the object of LDA is to use the covariances and means to define a line or *separatrix*, splitting the plane into two parts where observations will be assigned to either π_1 or π_2 , with the minimum cost (Methods). For higher dimensional data the separatrix is a hyperplane. QDA encompasses the case where the population covariances are unequal, $S_1 \neq S_2$, and the separatrix is a paraboloidal surface. As the name suggests QDA expands on LDA by including coefficients estimated for all quadratic interactions between the components of the observations (Eq. 3, Methods).

{Figure 10 about here please}

A measure that has gained some interest in characterising time series and $1/f^a$ noise is the Allan Variance (Allan 1966). The Allan Variance has been shown to be equivalent to Wavelet Variance for certain wavelets (Howe & Percival 1995) where the interest is in examining correlations at different scales. A further possible advantage of Allan Variance over the conventional variance is that it converges to a finite value for most types of noise, whereas the classical variance does not always converge to a finite value. For example the central-limit theorem does not apply to some chaotic processes (Yamaguchi & Nagai 1998) and the Allan Variance may be useful in such cases. The conventional Allan Variance is just the half the sum of the squared differences between adjacent points in a one dimensional series of numbers x_i : $AV = \frac{1}{2} \sum (x_{i+1} - x_i)^2 / N$. The factor of $\frac{1}{2}$ makes the Allan Variance equal to the conventional variance: $\sum (u - x_i)^2 / N$, (u is the population mean), for Gaussian distributed noise. Note that the AV is composed of the sum of squares of paired linear combinations of inputs of the form $(a-b)^2$. Such combinations represent a neurally plausible means (Discussion, Figure 17) of computing quadratic interactions between receptive field outputs, first proposed for motion detection (Emerson, Bergen & Adelson 1992).

In the section on mini-textures above we introduced the idea that the operation of spatial averaging *via* linear receptive field mechanisms would reduce the alphabet of mini-textures available for discrimination of the parent textures. Given that the Allan Variance, or some related measure, might be useful for characterising odd types of noise, and given the need to consider receptive field mechanisms we decided to examine a few variants of the Allan Variance in the receptive field outputs as inputs to our discriminant analyses. To do this we introduced square “receptive fields”, r , of various pixel sizes and spatial weightings. We then performed an operation akin to a convolution: calculating the mean output for the product of the receptive field weights with each pixel off the corresponding sized mini-texture m for positions i : $m_i r_i$. That is to say for each mini-texture the corresponding mini-texture pixel values and receptive field weights were multiplied point-wise, and their sum was divided by the number of pixels. The calculation differed from a convolution in that we only considered adjacent, non-overlapping, mini-textures. We then calculated the Allan Variance in these receptive field outputs for $AV = \frac{1}{2} \Sigma (\langle m_{i+1} r_{i+1} \rangle - \langle m_i r_i \rangle)^2 / N$. In fact the adjacent mini-textures for the whole sample texture were concatenated into one long row for this calculation so a few non-adjacent mini-textures were included into the calculation of AV. This was somewhat sloppy but the AV is a one-dimensional measure and the parent texture size was up to 256 pixels², which for small mini-textures meant that only a small number of these edge effects occurred. Nevertheless, we will later introduce some two dimensional variants of AV to resolve this problem and to address some other issues. We have included the one dimensional AV data for comparative purposes. In the analysis that follows we took the square root of AV to give the Allan Deviation, AD.

{Figure 11 about here please}

A comparison of AD and the conventional standard deviation (STD) showed that, at least for receptive fields smaller than 8 pixels², there is a significant difference between these

measures ($P < 0.01$). The data were based on twelve 256 pixel^2 textures made with 12 different random number seeds. Similar results were obtained for unoriented receptive fields and for receptive fields oriented at -45° and 45° (see Figs. 11, 12). The total range of receptive field sizes for which AD and STD were calculated included receptive field sizes that were powers of two: 2 to 128 pixels/side, and also $\{3, 5, 6, 7, 10, 14, 15, 21\}$ pixels/side.

The input to the Discriminant analyses was 12 ADs computed for 12 textures for each mini-texture/receptive field size combination. There were thus 12 observations for each case, \mathbf{x}_i , and we could increase the dimension of the observations and the discriminant models by increasing the number of receptive field sizes and types included in the model (Methods). We computed Receiver Operator Curves (ROCs) for each texture type and model. Figure 10 illustrates an ROC for the case of a model having 4 receptive fields, oriented and unoriented, each at two scales, 3×3 and 4×4 pixels (see inset Fig. 12), where the task was Odd vs. Random. The model was thus 4 dimensional, each observation \mathbf{x}_i having 4 AD measures. Twelve such observations were used to estimate the coefficients for the LDA and QDA models summarised in Eq. 2 and 3. The ordinate of Figure 10 shows the Sensitivity (Sens) or the probability of correctly recognising an Odd texture as Odd based on the receptive field outputs. The Specificity (Spec) is the probability of reporting Random when the pattern was Random. The abscissa is labelled $1-\text{Specificity}$, this difference being otherwise known as the “false alarm” or “false positive” rate, *i.e.* the probability of reporting Odd when the texture was Random. Notice that to achieve higher sensitivities the system generates increasing false positive errors. This is the classical ROC case for a recognising a signal containing noise. If there is no difference in the cost of falsely reporting Odd or Rand a convenient way to summarise the ROC is to report the simultaneously highest sensitivity and specificity. These points are marked with * on the LDA and QDA ROCs of Figure 10. We report these

simultaneous sensitivity and specificities henceforth, or their average, since they may differ slightly for our discretely computed ROCs (Maddess *et al.* 1999).

{Figure 12 about here please}

The outcome for QDA models containing receptive fields of different sizes, but all having the same orientation, is shown in Figure 11. Notice that as the number of receptive field sizes is increased discrimination performance improves. Some models with oriented receptive field sizes around 12 to 16 pixels² produced uniformly good discrimination across all textures but in doing so did not mimic human performance (not shown). Nevertheless, even for models including five receptive field sizes, performance is less than perfect; however, some of the features of the human data are captured.

Figure 11 suggests an obvious way to improve performance: use several orientations. We could not build models with very large numbers of orientations as this would require a large number of observations in order to estimate the covariance matrices (Methods). Employing a large number of observations to build a complex model would also not mimic our test conditions where relatively few targets of as few as 4x4 pixels were presented. We therefore examined models using a few oriented and unoriented receptive fields to span the orientation domain. Interestingly, models with relatively few receptive fields performed quite similarly to humans, even in the magnitude of the probability of making a correct judgement (Fig. 12). Figure 12 demonstrates another important outcome, also illustrated in Figure 10, that *LDA models never performed well* even with large numbers of receptive field types included in the models.

Thus far we used the Allan Deviation as the input to the discriminant models. As mentioned this is not especially suited to two dimensional textures. We therefore decided to examine some two dimensional variants of the AD and also to compare all these models with the classical standard deviation in the receptive field outputs. To compute the standard

deviation (STD) we simply took all the receptive field outputs from a given texture example and computed the standard deviation without regard to receptive field location. Notice that STD thus includes both short and long range interactions.

Three local versions of the AD were also considered: an unoriented case (*UnO*) and two oriented cases (*Or1*, *Or2*). Here instead of simply taking the mean squared differences between receptive field outputs along a line of the parent texture we took instead *weighted 2D differences*, squared these and took their mean. Thus, the difference in $\frac{1}{2} \sum (\langle m_{i+1}r_{i+1} \rangle - \langle m_i r_i \rangle)^2 / N$ is replaced by a local 2D differencing operation $D(\langle m_i r_i \rangle)$ to give $\sum D(\langle m_i r_i \rangle)^2 / N$. The weights D used for the three variants were:

<i>UnO</i>			<i>Or1</i>			<i>Or2</i>		
-1/8	-1/8	-1/8	-0.5	-0.5	1.0	1.0	-0.5	-0.5
-1/8	1.0	-1/8	-0.5	1.0	-0.5	-0.5	1.0	-0.5
-1/8	-1/8	-1/8	1.0	-0.5	-0.5	-0.5	-0.5	1.0

Thus, to compute each D , a set of 9 spatially adjacent $\langle m_i r_i \rangle$ were summed with the weights shown and the result was squared.

Since there appeared to be some similarity between the human data and the model outputs (*cf.* Figs. 5 to 7, and Fig. 12) we next attempted to quantitatively compare human data with the outputs from the various models. We first did this by taking the psychometric functions averaged across 5 subjects for ER, OR and EO comparisons of 32 pixel² patterns (Fig. 5) and we regressed these against means of the simultaneously maximum sensitivities and specificities obtained from versions of QDA models having different variance models. The regression took two forms, fitting a mean effect and a scaling, or only a scaling. The simpler single parameter scaling operation proved the best in terms of mean square errors and so we report that here, although the results were similar for the two parameter regression. We applied two further variations: fitting data and model outputs for all 9 gliders, or only data and model output for gliders 3 to 9. We will refer to these two cases as G_{all} and G_{39} . The G_{39}

method eliminated the possible effects of saturation of performance for patterns generated from gliders 1 and 2.

{Figure 13 about here please}

Figure 13 shows average QDA model output for 7 sets of 60 pixel² Even and Random input textures (circles, error bars) and the corresponding best fitting scaled psychometric functions for the case for ER comparisons (solid lines). In this case the model using the global variance (Var) in receptive field outputs provides data most closely mimicking the psychometric data for both G_{all} and G_{39} . The model data shown in the left and right columns differ in the tilt of the two oriented receptive fields (titles, caption) which were otherwise as shown in the inset Figure 12. Each panel of Figure 13 contains a legend describing the variance model, the t-statistic for the regression, and the resulting scaling factor required for the best match between model and human data. The left column shows fitted psychometric functions when data for all textures are compared (G_{all}), while the right column shows the outcome when only data for textures 3 to 9 are included in the regressions (G_{39}). Neither, G_{all} or G_{39} models produced significantly larger t-statistics on average. Of the 24 cases examined (G_{all} and G_{39} , using ER, OR, EO and Ave comparisons, for 60 120 and 240 pixel² input textures) the scaling factors were only significantly different than 1 in one case: EO for 240 pixel² inputs for the case of all 9 inputs, where the scaling was 0.86 ± 0.014 SE (P= 0.011). Figure 14 summarises the outcomes for all 24 cases. We also redid this analysis regressing against the psychometric functions obtained from 5 subjects for the 16 pixel² textures (Fig 5c). This produced very similar results as did using psychometric functions from 8 pixel² textures (although the scaling factors became larger).

{Figure 14 about here please}

We repeated the analysis of Figure 14, where the fits were based on data of Figure 5a, using both the psychometric functions obtained from the five repeats by the authors (Fig. 6a),

and for the combined psychometric functions obtained for different contrasts and pixel sizes (Fig. 7a). The mean t-statistics are summarised in Figure 15. If we want to examine the shape of the average functions in Figure 15 there is a problem in that different cases clearly have different mean t-statistics. A reasonable approach therefore is to fit a regression model estimating coefficients across the five variance models, also providing an offset for each case. In this way we extract the average function shape without injecting variance due to the different means for different cases. Overall the fit of these multiple regression models was good for both the G_{all} ($F=21.6_{df=16,44}$, $P < 0.0000$, $r^2 = 0.88$) and G_{39} ($F=39.9_{df=16,44}$, $P < 0.0000$, $r^2 = 0.93$) cases. The resulting regressed means are shown in Figure 15g, h.

We further examined models where a single mean was used for variance models Var and OR1 and these were not significantly ($P=0.16$) different to the case of separate means both G_{all} and G_{39} . Thus, the Var and OR1 variance models are about equally good matches to the psychometric data. All simpler models produced significantly poorer fits. For example using a single mean for All, Var and OR1, but separate means for UnO and Or2, was a significantly poorer model than that combining only Var and OR1 ($P= 0.017$ and 0.012 for the G_{all} and G_{39} cases). Thus, the most parsimonious regression model is one with a combined mean for Var and OR1 and different means for the other variance measures.

{Figure 15 about here please}

Discussion

A considerable volume of research examining the responses of single striate cortical neurons has concentrated on aspects of their spatial frequency and orientation tuning. The Fourier transform of the power spectrum of the spatial frequency tuning is the spatial autocorrelation function. Therefore these studies have in effect been concentrating on the two-point correlation properties of these cells. More recent work (Purpura *et al.* 1994), using isotrigon stimuli, has shown that almost all cells within primate V1 encode higher order correlations. Three point (third order) correlations, and higher order, inform us about spatial phase relationships, critical to the recovery of salient image components like edges. The results from V1 are supported by PET and fMRI data showing cortical areas apparently dedicated to the higher order correlation properties of visual stimuli (Beason-Held *et al.* 1998; Beason-Held *et al.* 1998; Beason-Held *et al.* 2000). Given the potential interest in isotrigon stimuli then we thought it would be useful to quantify for the first time the complete range of ER, OR and EO discriminations for the isotrigon stimuli used in psychophysical and VEP studies to date (Victor & Conte 1991). We also manipulated stimulus extent, contrast and pixel size for all these stimuli (Fig. 5, 6, 7, 8). Given the quite local analysis of cells in V1 we also sought to quantify the number and frequency of mini-textures within the 18 isotrigon stimuli examined (Tables 1, 2, 3, 1A). These findings now permit studies that compare isotrigon textures sharing no, or some number, of mini-textures. Our data also reveal that discrimination of some textures shows significant long-term learning effects (Fig. 9).

{Figure 16 about here please}

In a VEP study using Even textures of the Box type Victor and Conte (1989) found that spatial interactions critical to making ER discriminations proceed on scales proportional to pixel block size, at least for pixel blocks ranging from of 4' to 16' square, where this critical interaction distance is about 3 to 4 pixels. We confirm that psychometric

discrimination functions for all 9 glider types differ little for block pixel sizes 14.5' and 58' square (*cf.* Figs. 5, 6 with 7), and that this holds for OR and EO comparisons also. A comparative VEP and psychophysical study (Victor & Conte 1991) also examined Even *vs.* Random comparisons for the 9 glider types employed here, but where the sort-range correlations of the Even textures was manipulated. They characterised their data by fitting a model based on the rate of activation of hypothetical neural units (Victor & Conte 1990). In their model the amount of decorrelation of Even textures required for subjects to achieve 75% discrimination from Random binary textures is proportional to $1/\text{rate}$, a sort of delay to threshold. They tabulated their fitted rates for 6 subjects, and 9 Even (and 1 Odd) textures. Since the subjects had quite different mean rates we have fitted a multiple regression model similar to that for Figures 15g,h. Significantly different subject means were found and the overall model ($F=34.0_{df=14, 40}$, $P < 0.0000$) accounted for 91.7% of the variance. To provide reasonably Gaussian distributed data we fitted the square root of the rates (Fig. 16a), and later squared these values to produce Figure. 16b. Both figures show similarities to our measured ER discrimination functions (*cf.* Figs 5 to 8), particularly in having dips at glider/texture numbers 3 (Cross), 7 (Wye) and 9 (EI).

In the analysis summarised in Figures 13 to 15 we directly compared the probability of correct discrimination by our subjects with the mean of simultaneously highest sensitivity and specificity of the QDA models. There is no *a priori* reason why these probabilities should be the same size, nevertheless the required scaling factors only differed from 1 of the 60 cases examined for comparison with psychophysical data for 32 pixel² textures. As noted in the Results the reason for using the simultaneously highest sensitivity and specificity is that it represents the best that a model governed by a ROC can do when the penalty for incorrect decisions is the same for either type of test pattern, as was the case for our subjects. Thus, the fact that the scaling factors differed little from one, indicated that humans behaved like an

ideal observer using a discrimination process based on measures related to those employed in our models.

Models of texture segregation have concentrated on identifying the basic mechanisms that must underlie this ability (Beck *et al.* 1987; Sutter *et al.* 1989; Landy & Bergen 1991; Graham *et al.* 1992; Graham & Sutter 1996; Graham & Sutter 2000). The present study seeks to compliment these analyses by introducing both isotrigon textures and formal statistical discriminant models. Part of the concept behind the discriminant models is to compare quadratic and linear models given a fairly limited number of samples, as the human subjects had in our tests. A secondary objective was to see if either type of discriminant model would show behaviour that was anything like that of human subjects.

The host of data indicating initial processing of textures by neurons with small, linear, oriented receptive fields, suggested that examining small texture domains, mini-textures, might be sensible. Also, linear summation by these receptive fields would clearly reduce the alphabet of mini-textures available for discrimination. Indeed, a simple model indicated that discrimination is not directly related to number of mini-textures available (Fig. 8b). Nevertheless, the number of mini-textures unique to given texture types was not a very poor predictor of discrimination either, suggesting that some modification of mini-texture number, as spatial summation would afford, determined discrimination performance.

While it is clear that linear summation acts to reduce the number of independent inputs possible from a given texture class, it is equally clear that that *nonlinear interactions can increase the alphabet size*. For example if one had two independent linear receptive field outputs, a and b , for two (or more) types of mini-texture, then $(a \pm b)^2$ leads to response terms in a^2 , ab and b^2 , three independent measures. If we consider products of linear combinations of a and b up to order d , then dimensionality of the *feature space* into which the discrimination problem is cast grows as 2^d , and N^d , for N inputs (Schölkopf, Smola &

Müller 1998) (see also (Maddess, Davey & Yang 1999)). Notice that the Allan Variance measure we selected has the form $(a-b)^2$, being based on the square of local differences (with range -1 to 1) of mean receptive field output in response to mini-texture input. As outlined in the Methods, the Allan Variance and its derivatives may have some utility compared to traditional variance measures, and is related to wavelet variance (Howe & Percival 1995).

None of the linear discriminant models we made gave good discrimination (*e.g.* Figs. 10 and 12). This is not entirely surprising because in order to discriminate an ensemble of isotrigon textures from another group by definition requires examination of fourth order correlations (Victor 1994). The linear comparison of variances afforded by LDA is formally second order. By contrast, the quadratic discrimination models (QDA) we examined performed more like humans, even in the detailed shape and magnitude of discrimination the discrimination functions (*e.g.* Figs. 12 and 13). Presumably the quadratic models work because, by being based on quadratic interaction between variance measures (Eq. 3), they are formally *fourth order*. A necessary condition for good performance was that the initial linear spatial filtering had to be done by receptive fields that spanned the orientation domain and which operate at (at least) a few, small, scales (*cf.* Figs. 12 and 13). Victor and Conte (1989) point out that operation $(w+x+y+z)^4$, where x to z are receptive field outputs, would yield products in $wxyz$, *i.e.* four point correlations. As indicated by those authors, however, the other resulting products, x^2wy *etc.* would make the $wxyz$ term relatively small, and a tiny fraction of any neuronal response. Nevertheless, the above discussion suggests that we might have obtained similar performance by raising our local differences to the fourth power: $(\langle m_{i+1}r_{i+1} \rangle - \langle m_i r_i \rangle)^4$, and then constructing linear discriminant models on that output. Psychophysical studies of putative complex channels in texture segregation indicate exponents between 3 and 4 are involved (Graham & Sutter 1998), but whether these are

generated by a single neural mechanism or represent cascaded, less steeply accelerating, responses is unclear.

A possible advantage of our formulation with two levels of quadratic interaction is that if adaptation is introduced at each stage (as for example in taking the mean rather than the summed receptive field output) then such a neural system would be *less likely to saturate*. Adaptation is a well-known feature of striate cortical neurons (*e.g.* (Maddess, McCourt, Blakeslee & Cunningham 1988; Carandini & Heeger 1994; Carandini, Heeger & Movshon 1997)).

{Figure 17 about here please}

Physiologically feasible models exist for obtaining squares of linear combinations of inputs and have been used to construct physiologically plausible versions of the Reichardt and Motion Energy models using only rectification and summation (Emerson *et al.* 1992). If one removes the delay stage from these models measures like the Allan Variance can be obtained (see below). Moreover, these models get rid of unwanted products (as occurs in the case of $(w+x+y+z)^4$ above), and so there may be less need for gain control to prevent saturation. So, for two pixels a, b two products are computed (Figure 17) and differenced: $(a+b)^2 - (a-b)^2 = 4ab$, thus eliminating the self quadratic terms in a^2 and b^2 . Cascading this type of operation, and introducing two other pixels u and v , we obtain: $(ab+uv)^2 - (ab-uv)^2 = 4abuv$, a pure four point correlation without other products to cause saturation problems for any neuron. If ab and uv are replaced with averages of similar products from nearby locations, $\langle a_i b_i \rangle$ and $\langle u_i v_i \rangle$, and a to v are differences from mean pixels outputs (contrasts) then we essentially have the models used here that are based on quadratic interactions between local variances. The present models, being cascades of two quadratic stages of processing, also resemble some models for non-Fourier motion detection (Zanker 1996).

Part of the reason for introducing the 2D versions of the Allan Variance used here was that previous work on discrimination isotrigon textures indicated that, for textures consisting of pixel blocks greater than 4' on a side, most of the important spatial interactions (at least for ER comparisons) took place over a range of about 3 pixels (Victor & Conte 1989). Our unoriented and oriented Allan Deviations (*UnO*, *OR1*, *OR2*), in combining information arising from mini-textures as large as 4 pixels², include interactions out to about 8 pixels from a central pixel. Nevertheless, the global variance measure, *Var*, provided as good a model as *Or1*, and better than *UnO* or *All*. This suggests long-range correlations may be more important than was thought. At the same time the case for accepting the *OR1* model is supported by texture segregation models where both input and output stages are oriented (Landy & Bergen 1991; Graham *et al.* 1992; Graham & Sutter 1996; Graham & Sutter 1998; Graham & Sutter 2000). A curious feature of the present models was that *OR1* was better than *OR2* whether or not the orientations of the input receptive fields and the subsequent differencing operator were parallel or orthogonal. Apparently the orientation of the latter operator was the important feature. Some recent texture segregation studies include mixtures of linear and nonlinear inputs to the (undefined) decision process, and or interactions between these units (Graham & Sutter 1998; Graham & Sutter 2000). Our analysis says nothing about such interactions but perhaps the discrimination of the Triangle texture might be facilitated by such a mechanism. Of all the textures tested it is only isodipole with random textures. Thus, the fourth order mechanisms described above are not required to discriminate Triangle textures from random textures.

A proposed model for ER discriminations (Victor & Conte 1991) comprised oriented rectifying subunits followed by a strong threshold. A key feature of this model is the summation of 6 or more of these rectifying units to make a rather elongated receptive field, which would have a very narrow orientation tuning. The model was tested only on ER

comparisons for the standard Box glider textures. In our case the use of larger oriented input receptive fields sometimes yielded good performance, but discrimination by these models was quite uniform across the different texture types, unlike human performance. Very elongated receptive fields are quite rare in striate cortex. Highly orientation tuned and rectifying neurons are reported in cat area 21a (Dreher, Michalski, Ho, Lee & Burke 1993), however, and also in bees (Yang & Maddess 1997), which have been shown capable of discriminating isotrigon textures (Maddess *et al.* 1999).

Whether our actual neural machinery relies upon a single rapidly accelerating nonlinearity (Graham & Sutter 1998) or a cascade of nonlinearities as used in the present paper and elsewhere (Victor & Conte 1991), it seems clear that consideration of higher order image correlations is necessary for understanding image structure and such information is encoded even by most striate cortical cells (Purpura *et al.* 1994). As pointed out above cascade models may have advantages from the point of view of reducing saturation problems. Certainly, nonlinear interactions seem to be related to responses to contours (Polat, Mizobe, Pettet, Kasamatsu & Norcia 1998). More generally, a number of sophisticated methods for separation of signals, including Independent Component Analysis and Nonlinear Principal Component Analysis, have all been shown to rely on fourth order correlations (Lee, Girolami, Bell & Sejnowski. 2000) to minimise mutual information in order to achieve segregation. It would be surprising if our brains did not use related methods to solve signal segregation and discrimination problems.

References

- Allan, D.W. (1966). Statistics of atomic frequency standard. *Proc. IEEE*, **54**, 221-231.
- Beason-Held, L.L., Purpura, K.P., Krasuski, J.S., Desmond, R.E., Mangot, D.J., Daly, E.M., Optican, L.M., Rapoport, S.I. & VanMeter, J.W. (2000). Striate cortex in humans demonstrates the relationship between activation and variations in visual form. *Exp Brain Res*, **130**, 221-6.
- Beason-Held, L.L., Purpura, K.P., Krasuski, J.S., Maisog, J.M., Daly, E.M., Mangot, D.J., Desmond, R.E., Optican, L.M., Schapiro, M.B. & VanMeter, J.W. (1998). Cortical regions involved in visual texture perception: a fMRI study. *Brain Res Cogn Brain Res*, **7**, 111-8.
- Beason-Held, L.L., Purpura, K.P., Van Meter, J.W., Azari, N.P., Mangot, D.J., Optican, L.M., Mentis, M.J., Alexander, G.E., Grady, C.L., Horwitz, B., Rapoport, S.I. & Schapiro, M.B. (1998). PET reveals occipitotemporal pathway activation during elementary form perception in humans. *Vis Neurosci*, **15**, 503-10.
- Beck, J. (1983). Textural segmentation, second-order statistics, and textural elements. *Biol Cybern*, **48**, 125-130.
- Beck, J., Sutter, A. & Ivry, R. (1987). Spatial frequency channels and perceptual grouping in texture segregation. *Comput. Vision Graphics Image Process.*, **37**, 299-325.
- Carandini, M. & Heeger, D.J. (1994). Summation and division by neurons in primate visual cortex. *Science*, **264**, 1333-6.
- Carandini, M., Heeger, D.J. & Movshon, J.A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci*, **17**, 8621-44.
- Chubb, C. & Sperling, G. (1991). Texture quilts: basic tools for studying motion-from-texture. *J. Math. Psychol.*, **35**, 411-442.

- Dreher, B., Michalski, A., Ho, R.H.T., Lee, C.W.F. & Burke, W. (1993). Processing of form and motion in area 21a of cat visual cortex. *Vis. Neurosci.*, **10**, 93-115.
- Emerson, R.C., Bergen, J.R. & Adelson, E.H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Res.*, **32**, 203-218.
- Gagalowicz, A. (1981). A new method for texture fields synthesis: some applications to the study of human vision. *IEEE Trans Pat. Anal. Mach. Int.*, **3**, 520-533.
- Gilbert, E.N. (1980). Random colorings of a lattice on squares in the plane. *SIAM J. Algebra Dis. Meth.*, **1**, 152-159.
- Graham, N. & Sutter, A. (1996). Effect of spatial scale and background luminance on the intensive and spatial nonlinearities in texture segregation. *Vision Res.*, **36**, 1371-1390.
- Graham, N. & Sutter, A. (1998). Spatial summation in simple (Fourier) and complex (non-Fourier) texture channels. *Vision Res.*, **38**, 231-57.
- Graham, N. & Sutter, A. (2000). Normalization: contrast-gain control in simple (Fourier) and complex (non-Fourier) pathways of pattern vision. *Vision Res.*, **40**, 2737-2762.
- Graham, N., Sutter, A., Venkatesan, C. & Humaran, M. (1992). Non-linear processes in perceived region segregation: orientation selectivity of complex channels. *Ophthalmic Physiol Opt.*, **12**, 142-146.
- Howe, D.A. & Percival, D.B. (1995). Wavelet variance, Allan variance, and leakage. *IEEE Trans. Inst. Measure.*, **44**, 94-97.
- Johnson, R.A. & Wichern, D.W. (1992). Applied multivariate statistical analysis. New Jersey, Prentice Hall.
- Julesz, B., Gilbert, E.N., Shep, L.A. & Frisch, H.L. (1973). Inability of humans to discriminate between visual textures that agree in second-order statistics - revisited. *Perception*, **2**, 391-405.

- Julesz, B., Gilbert, E.N. & Victor, J.D. (1978). Visual discrimination of textures with identical third-order statistics. *Biol. Cybern.*, **31**, 137-140.
- Kingdom, F.A., Keeble, D. & Moulden, B. (1995). Sensitivity to orientation modulation in micropattern-based textures. *Vision Res*, **35**, 79-91.
- Klein, S.A. & Tyler, C.W. (1986). Phase discrimination of compound gratings: generalized autocorrelation analysis. *J Opt Soc Am A*, **3**, 868-879.
- Landy, M.S. & Bergen, J.R. (1991). Texture segregation and orientation gradient. *Vision Res*, **31**, 679-691.
- Lee, T.-W., Girolami, M., Bell, A.J. & Sejnowski, T.J. (2000). A Unifying Information-theoretic Framework for Independent Component Analysis. *Computers math. app.*, **31**, 1-21.
- Maddess, T., Davey, M. & Yang, E. (1999). Discrimination of complex textures by bees. *J. Comp. Physiol. A*, **184**, 107-177.
- Maddess, T., Goldberg, I., Wine, S., Dobinson, J., Welsh, A.H. & James, A.C. (1999). Testing for glaucoma with the spatial frequency doubling illusion. *Vision Res.*, **39**, 4258-4273.
- Maddess, T., James, A.C., Goldberg, I., Wine, S. & Dobinson, J. (2000). A frequency doubling illusion based multiregion PERG for glaucoma. *Invest. Ophthalmol. Vis. Sci.*, **41**, 3818-3826.
- Maddess, T., McCourt, M.E., Blakeslee, B. & Cunningham, R.B. (1988). Factors governing the adaptation of cells in area-17 of the cat visual cortex. *Biol. Cybern.*, **59**, 229-236.
- Polat, U., Mizobe, K., Pettet, M.W., Kasamatsu, T. & Norcia, A.M. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature*, **391**, 580-4.

- Purpura, K.P., Victor, J.D. & Katz, E. (1994). Striate cortex extracts higher-order spatial correlations from visual textures. *Proc Natl Acad Sci U S A*, **91**, 8482-8486.
- Schölkopf, B., Smola, A. & Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comp.*, **10**, 1299 - 1319.
- Sutter, A., Beck, J. & Graham, N. (1989). Contrast and spatial variables in texture segregation: testing a simple spatial-frequency channels model. *Percept Psychophys*, **46**, 312-332.
- Turano, K. & Pantle, A. (1989). On the mechanism that encodes the movement of contrast variations: velocity discrimination. *Vision Res*, **29**, 207-221.
- Victor, J.D. (1985). Complex visual textures as a tool for studying the VEP. *Vision Res.*, **25**, 1811-1827.
- Victor, J.D. (1994). Images, statistics and textures: implications of triple correlation uniqueness for texture statistics and the Julesz conjecture: comment. *J. Opt. Soc. Am.*, **A11**, 1680-1684.
- Victor, J.D. & Conte, M.M. (1989). Cortical interactions in texture processing: scale and dynamics. *Vis. Neurosci.*, **2**, 297-313.
- Victor, J.D. & Conte, M.M. (1990). Motion mechanisms have only limited access to form information. *Vision Res.*, **30**, 289-301.
- Victor, J.D. & Conte, M.M. (1991). Spatial organization of nonlinear interactions in form perception. *Vision Res.*, **31**, 1457-1488.
- Victor, J.D. & Zemon, V. (1985). The human visual evoked potential: analysis of components due to elementary and complex aspects of form. *Vision Res.*, **25**, 1829-1842.
- Werkhoven, P., Sperling, G. & Chubb, C. (1993). The dimensionality of texture-defined motion: a single channel theory. *Vision Res*, **33**, 463-485.

- Werkhoven, P., Sperling, G. & Chubb, C. (1994). Perception of apparent motion between dissimilar gratings: spatiotemporal properties. *Vision Res*, **34**, 2741-2759.
- Yamaguchi, H. & Nagai, Y. (1998). Correlation dimension as the large deviation for distribution of distances between two points on an embedded manifold. *J. Phys. Soc. Japan*, **10**, 3397-3404.
- Yang, E.C. & Maddess, T. (1997). Orientation-sensitive neurons in the brain of the honeybee (*Apis mellifera*). *J Insect Physiol*, **43**, 329-336.
- Yellott, J.I. (1993). Implications of triple correlation uniqueness for texture statistics and the Julesz conjecture. *J. Opt. Soc. Am. A*, **10**, 777-793.
- Zanker, J.M. (1996). On the elementary mechanism underlying secondary motion processing. *Philos Trans R Soc Lond B Biol Sci*, **351**, 1725-36.

Tables

Table 1

No.	Name	3x3	4x4
1	Box	32	128
2	Triangle	32	128
3	Cross	256	4096
4	Zigzag	128	1024
5	Oblong	128	1024
6	Tee	128	1024
7	Wye	256	4096
8	Foot	128	1024
9	El	128	1024

Table 2

(a)

E^E	Box	Tri	Cro	Zig	Obl	Tee	Wye	Foo	El
Box	32	4	16	16	32	8	16	16	8
Tri	4	32	16	8	8	8	16	32	8
Cro	16	16	256	64	64	64	128	64	64
Zig	16	8	64	128	32	32	64	32	32
Obl	32	8	64	32	128	32	64	32	32
Tee	8	8	64	32	32	128	64	32	32
Wye	16	16	128	64	64	64	256	64	64
Foo	16	32	64	32	32	32	64	128	32
El	8	8	64	32	32	32	64	32	128

(b)

O^O	Box	Tri	Cro	Zig	Obl	Tee	Wye	Foo	El
Box	32	0	16	16	0	8	16	0	8
Tri	0	32	16	8	8	8	16	0	8
Cro	16	16	256	64	64	64	128	64	64
Zig	16	8	64	128	32	32	64	32	32
Obl	0	8	64	32	128	32	64	32	32
Tee	8	8	64	32	32	128	64	32	32
Wye	16	16	128	64	64	64	256	64	64
Foo	0	0	64	32	32	32	64	128	32
El	8	8	64	32	32	32	64	32	128

(c)

E^O	Box	Tri	Cro	Zig	Obl	Tee	Wye	Foo	El
Box	0	4	16	16	0	8	16	16	8
Tri	0	0	16	8	8	8	16	0	8
Cro	16	16	0	64	64	64	128	64	64
Zig	16	8	64	0	32	32	64	32	32
Obl	32	8	64	32	0	32	64	32	32
Tee	8	8	64	32	32	0	64	32	32
Wye	16	16	128	64	64	64	0	64	64
Foo	0	32	64	32	32	32	64	0	32
El	8	8	64	32	32	32	64	32	0

Table 3

No.	Name	Mean N Even	SD N Even	Mean N Odd	SD N Odd
1	Box	300.1 ± 22.7	82.9 ± 29.8	300.1 ± 15.3	48.6 ± 8.7
2	Triangle	300.1 ± 2.6	15.8 ± 6.6	300.1 ± 2.9	15.8 ± 5.9
3	Cross	37.5 ± 3.0	11.5 ± 3.2	37.5 ± 2.5	8.9 ± 2.0
4	Zigzag	75.0 ± 6.9	27.2 ± 10.7	75.0 ± 5.4	18.0 ± 7.9
5	Oblong	75.0 ± 8.4	26.0 ± 7.9	75.0 ± 3.7	15.6 ± 3.1
6	Tee	75.0 ± 2.2	8.7 ± 2.5	75.0 ± 2.2	8.2 ± 2.5
7	Wye	37.5 ± 1.9	6.2 ± 1.7	37.5 ± 1.3	6.0 ± 1.2
8	Foot	75.0 ± 3.0	16.3 ± 5.8	75.0 ± 2.7	10.6 ± 1.8
9	El	75.0 ± 2.2	8.4 ± 2.7	75.0 ± 2.8	8.7 ± 2.6

Table A1

Glider	Glider Rule	N(D(N,M))	N _{3x3}
Box	$S(i+1,j+1) = F(S(i,j), S(i+1,j), S(i,j+1))$	2^{N+M-1}	$2^5=32$
Triangle	$S(i+1,j+1)=F(S(i+1,j),S(i,j+1))$	2^{N+M-1}	$2^5=32$
Cross	$S(i+1,j+2)=F(S(i,j+1),S(i+1,j),S(i+2,j+1))$	$2^{2N+2M-4}$	$2^8=256$
Zigzag	$S(i+1,j+1)=F(S(i+1,j),S(i+2,j),S(i,j+1))$	2^{N+2M-2}	$2^7=128$
Oblong	$S(i+2,j+1)=F(S(i,j),S(i+2,j),S(i,j+1))$	2^{N+2M-2}	$2^7=128$
Tee	$S(i+2,j+1)=F(S(i+1,j),S(i,j+1),S(i+1,j+1))$	2^{N+2M-2}	$2^7=128$
Wye	$S(i+2,j+2)=F(S(i+1,j),S(i,j+1),S(i+1,j+1))$	$2^{2N+2M-4}$	$2^8=256$
Foot	$S(i+2,j+1)=F(S(i+1,j),S(i+2,j),S(i,j+1))$	2^{N+2M-2}	$2^7=128$
El	$S(i+2,j+1)=F(S(i,j),S(i+1,j),S(i,2,j))$	2^{N+2M-2}	$2^7=128$

Table Legends

Table 1. The number of mini-textures found for 3x3 and 4x4 pixel domains in each of the nine texture types using the histogram method of Fig. 3. The number of mini-textures for Odd and Even types of each texture is the same but none of the mini-textures are shared between Odds and Evens (see Table 2c). Data were reproduced for textures generated from 16 different random number seeds and the same mini-textures were observed for each texture type independent of the seeds.

Table 2. The number of *shared* 3x3 mini-textures for the 9 texture types. The labels are the first 3 letters of the glider names from Table 1. **(a)** Number of shared 3x3 textures between Even textures. The diagonal (bold) is the same as column 3 of Table 1. **(b)** Shared mini-textures between Odd textures. **(c)** Shared mini-textures between Even and Odd. Notice that the diagonal is all zeros (for an example see Fig. 4).

Table 3. The mean frequency of occurrence [number / (100 x 100 pixel texture)] of 3x3 mini-textures for Even (column1) and Odd (column 3) versions of the 9 texture types. The standard deviation in those quantities are also given (columns 2 and 4). Each value in the table has an associated standard deviation (\pm SD) since they are each an estimate from 16 repeats for textures generated from initial data sets based on 16 different random number seeds.

Figure Legends

Figure 1. Examples of the texture types used. The 18 texture types illustrated in the bottom two rows were produced using a recursion rule (Methods). At the top of each column is a representation of the *glider* used to produce the Even and Odd textures below it. The white block pixels represent the active pixels of the glider (Methods). Even and Odd textures (rows 2 and 3) being produced by two variants of the recursion rule operated on each glider type. The names of the gliders, and so their related texture types, are shown above along with a numerical reference {1 to 9} used in place of the glider names in figures that follow.

Figure 2. The temporal function governing the contrast of the textures during each presentation. The sigmoidal onset and offset are determined by a Blackman function (Methods) and were introduced to reduce nonlinearities associated with abrupt changes in contrast.

Figure 3. Illustration of the mini-textures found in Even (a), and Odd (b), E1 textures (texture type 9, Fig. 1). The 34,848 3x3 pixel mini-textures contained in each of 32 one hundred pixel² textures generated using 32 different random number seeds were sorted into the 512 possible mini-texture types permitted for 3x3 binary mini-textures. The histogram (c) shows accumulated mini-texture counts for Odd (black bars) and Even (grey bars) versions of E1 textures. Each generates 128, mutually exclusive, mini-texture types. (d) The frequency of each mini-texture type per 100 pixel² texture for Even (solid line) and Odd (dashed line) versions. The mean frequency + 1 SD for Even (dotted line), and the mean frequency - 1 SD for Odd (dash dot line) are also shown.

Figure 4. The 32 mini-textures found in Even (upper 4 rows) and Odd textures (lower 4 rows) produced with the Box glider. Notice that each Even mini-texture has a corresponding Odd mini-texture that differs by one pixel.

Figure 5. Average psychometric functions showing the mean probability (P) of correct discriminations from 5 subjects for ER, OR and EO (a, legend) discriminations (a,c,d) and the averages for three texture sizes types (b,d,f). The top row of panels (a,b) shows data for 32 pixel² textures, the middle (c,d), and bottom rows (e,f) for 16 and 8 pixels² respectively. Texture numbers shown on the abscissa are defined in Fig. 1. Error bars are SE and one set per panel are shown to maximise legibility.

Figure 6. As for Fig. 5 but the data are averages of from 5 repeats by each of the authors (10 data sets in all). Different seeds for the random number generator (Methods) were used for each of the 5 tested sets of textures. Conventions as in Fig. 5.

Figure 7. Mean probability of correct discrimination for different contrasts and block pixel size (a, legend). Each curve in 7a,c,e is the average of data for the two authors and their ER, OR and EO discriminations. Conventions otherwise as in Fig. 5.

Figure 8. Mean probability of correct discriminations for 4 pixel² textures. (a) Each curve is the average of 2 data sets collected during the first two repeats of the data shown in Fig. 6. (b) The average of the three curves of (a, dashed line) with SE. The solid line is a simple model based upon the number of mini-textures per texture (Results).

Figure 9. The fitted slopes of the probability/trial over the 5 trials of Fig. 6. A positive slope indicates the increase in the probability correct (Δp) across the 5 trials. The three lines represent outcomes for square textures having the side lengths shown in the legend. For each texture two models were fit, the simpler had two parameters: a mean probability and a slope, while the second model had separate means for the two subjects and a slope. In each case F-tests determined if the more complex, 3 parameter, model was justified at the $p < 0.05$ level. Circles indicate a significant slope ($p < 0.05$), * indicate the 3 parameter model was justified.

Figure 10. Example of Receiver Operator Curves (ROCs) for linear (+, LDA) and quadratic (o, QDA) discriminant models. The particular case is for Cross textures and an OR

comparison. The QDA curve is one of the 9 ROCs computed to generate Fig. 12. Notice that as the probability of detecting Odd patterns from binary noise when presented with an Odd (Sensitivity) increases, the probability of falsely reporting a Random pattern as being Odd (1-Specificity) also increases. If there is no difference in the cost of falsely reporting Odd or Random then a good summary parameter for the ROC is the maximum simultaneously highest sensitivity and specificity (*'s). Notice that the LDA model performs worse than the QDA model.

Figure 11. Sensitivity and specificity for QDA models using combinations of receptive field (insets top right); and different sizes and numbers (legend). The simultaneously highest sensitivity and specificities are averaged together because, while they are theoretically the same, they occasionally differ due to the numerical procedure used to generate the ROCs (see Fig. 10). The example is for ER comparisons and one receptive field orientation: unoriented (top panel), 45° (middle panel), -45° (bottom) panel (*e.g.* insets). Discrimination improves as the number of receptive fields of different sizes increases. While the curves capture some of the qualities of the human psychometric functions they tend to have one or more major departures (especially for textures 1,2 and 6), some receptive field orientations apparently being more useful than others for detecting particular textures.

Figure 12. Summary of a 4 receptive field model with mixed orientations (inset, lower). The ordinate is the simultaneously highest sensitivity (Sens, circles) and specificity (Spec, squares) (see Fig. 10). (Upper) Linear discriminant models fail to get much above 75% correct in any case. The specific example is for OR comparison. (Lower) the quadratic discriminant models show better overall performance except for textures 7 and 9, thus closely mimicking human performance (*cf.* Figs. 5, to 8).

Figure 13. Comparison of QDA output and human data. All QDA models were like that for Fig. 12 using 3x3 and 4x4 pixel receptive fields: 2 oriented and 2 unoriented. In this particular

example all panels are ER comparisons and the solid curves are scaled versions of the 32 pixel^2 psychometric functions for ER comparisons for 5 subjects from Fig. 5a. The *left column* (a, c, e, g, i) is for models where the receptive fields were unoriented or -45° and where the regression included data for textures 1 to 9. The *right column* (b, d, f, h, j) is for models with receptive fields that were unoriented or 45° , and where the regression included only model and human data from textures 3 to 9. The 5 *rows* of panels summarise results for QDA models where different variance models were used. **From top to bottom** the models were: *All* Allan Deviation, *UnO* unoriented 2D Allan Deviation, *Var* global Standard Deviation, *Or1* -45° oriented 2D Allan Deviation, *Or2* 45° oriented 2D Allan Deviation. Each case the t-statistic is reported for a one parameter regression model providing a best scaling of the psychometric function for to the particular model output. The scale value, b_i is also given. *All the scale values are close to 1*. The t and b values in both columns are for the right column cases only.

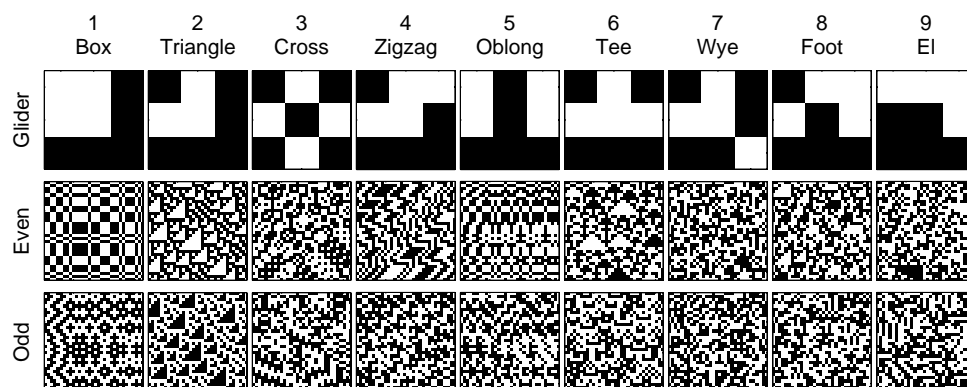
Figure 14. The process illustrated in Fig. 13 was repeated for ER, OR and OE comparisons (legends) and for 3 texture sizes, 60 (a, b), 120 (c, d), and 240 pixels^2 (e, f). Seven examples of each texture type were generated for each texture size using different random number seeds. In this way the five variance models (abscissa labels defined in text) were compared for smaller to larger mini-texture sample sizes. Larger training sets did not perform better. Each panel shows the t values obtained for the best scaling of the psychophysical data for different comparisons. In the case of the curves labelled *Ave* the fits were between the averaged psychometric function and the average of the model discriminant functions for ER, OR and EO. The horizontal dashed line indicates the t value for a nonsignificant regression. The bottom row (g, h) represents the means across the three texture sizes. The left column of panels (a, c, e, g) are t values when regressing data for all 9 textures, while the right column

(b, d, f, h) is for regressing model and data for textures 3 to 9 only. Except for g, h (see Fig. 15) the SE (N=7) are about the size of the symbols.

Figure 15. The modelling process summarised in Figs. 13 and 14 was repeated using the psychometric functions derived from the data sets illustrated in Figs. 5 to 7. As in Fig. 14 columns of panels represent regressing data for all 9 gliders (a, c, e) and only gliders 3 to 9 (b, d, f). Each panel is a mean for 3 texture sizes as in Fig. 14g, h, which are reproduced as c,d here. The mean shape of the curves in **a to f** as determined by a multiple regression model are shown in g, h. One set representative error bars (SE, N=3) for each panel are shown.

Figure 16. (a) The mean square roots of the *rates* tabulated by Victor and Conte (Victor & Conte 1991). $1/\text{rate}$ is a measure of the delay in time to threshold (75% performance) for a hypothetical set of neurons. Larger rates thus indicate the degree of local decorrelation that could be introduced into Even versions of the textures used before discrimination from random textures failed. **(b)** The plot shows $0 - 1/\text{rate}$, where rate is derived from the fitted values shown in (a). The vertical inversion was done to have downward displacement indicate poorer performance.

Figure 17. A physiologically plausible model for quadratic interactions using rectification and summation. **(top)** Neuron 1 having a *soft* (quadratic-like) half-wave rectified response to the summed input from two pixels A and B. The summation $A+B$ could occur by both pixels falling within a positively weighted receptive field. **(middle)** Neuron 2 responding similarly but to positive excursions of $A+B$. **(bottom)** A third neuron sums the outputs, R and S, from the first two neurons. Note that due to the soft rectification $R+S \approx (A+B)^2$. Differences described in the text, like $A-B$, can be computed by receptive fields like the partial derivative (d/dx) of a 2D Gaussian, with A and B placed in the positive and negative lobes respectively, providing $(A-B)^2$. A subsequent difference between the two squares, $(A+B)^2 - (A-B)^2$, gives $4AB$. Cascading two similar processes can provide pure fourth order correlations.

**Figure 1**

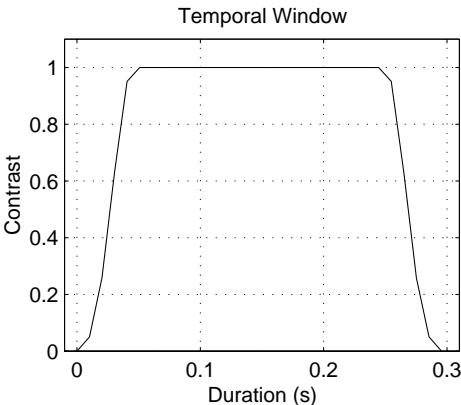
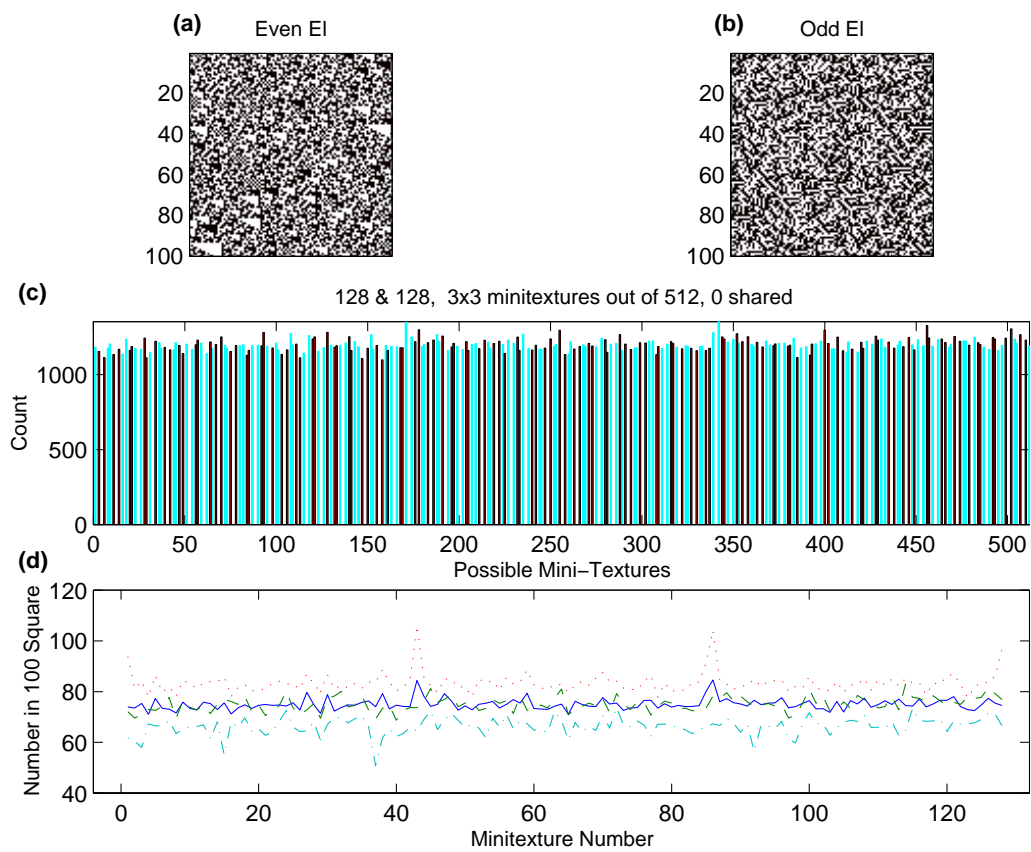


Figure 2

**Figure 3**

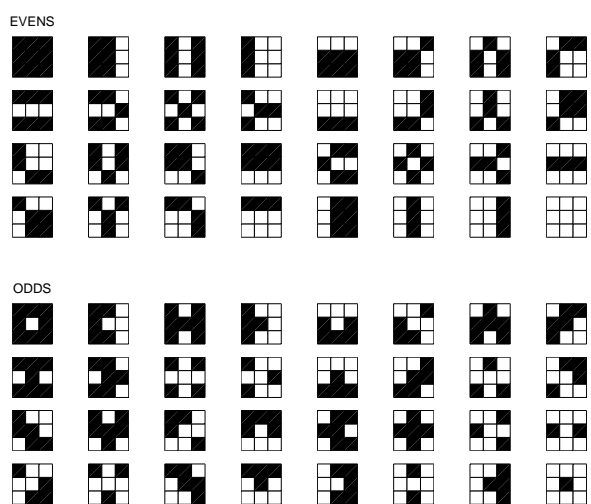


Figure 4

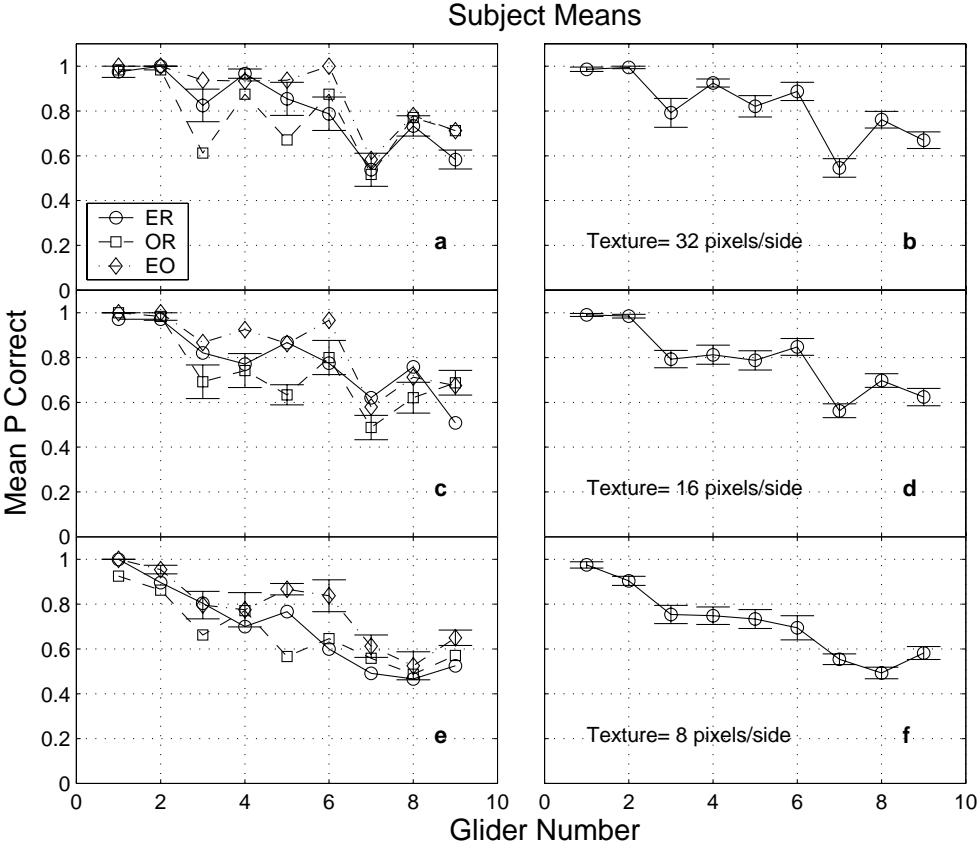


Figure 5

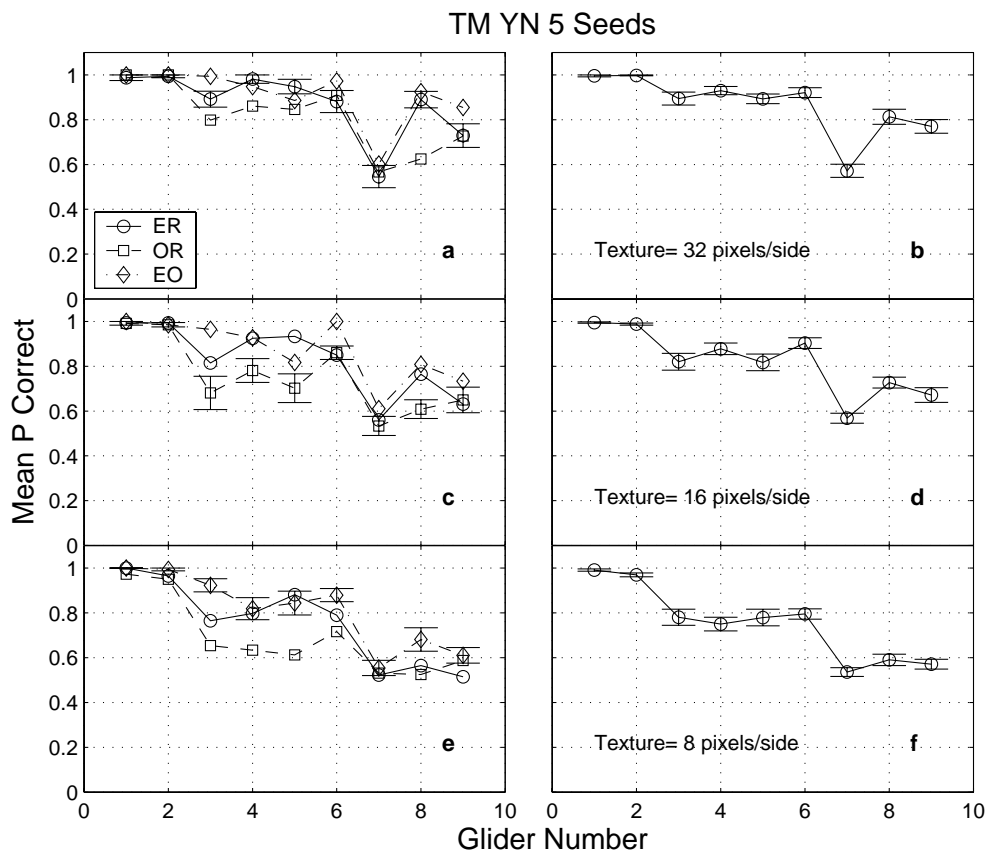


Figure 6

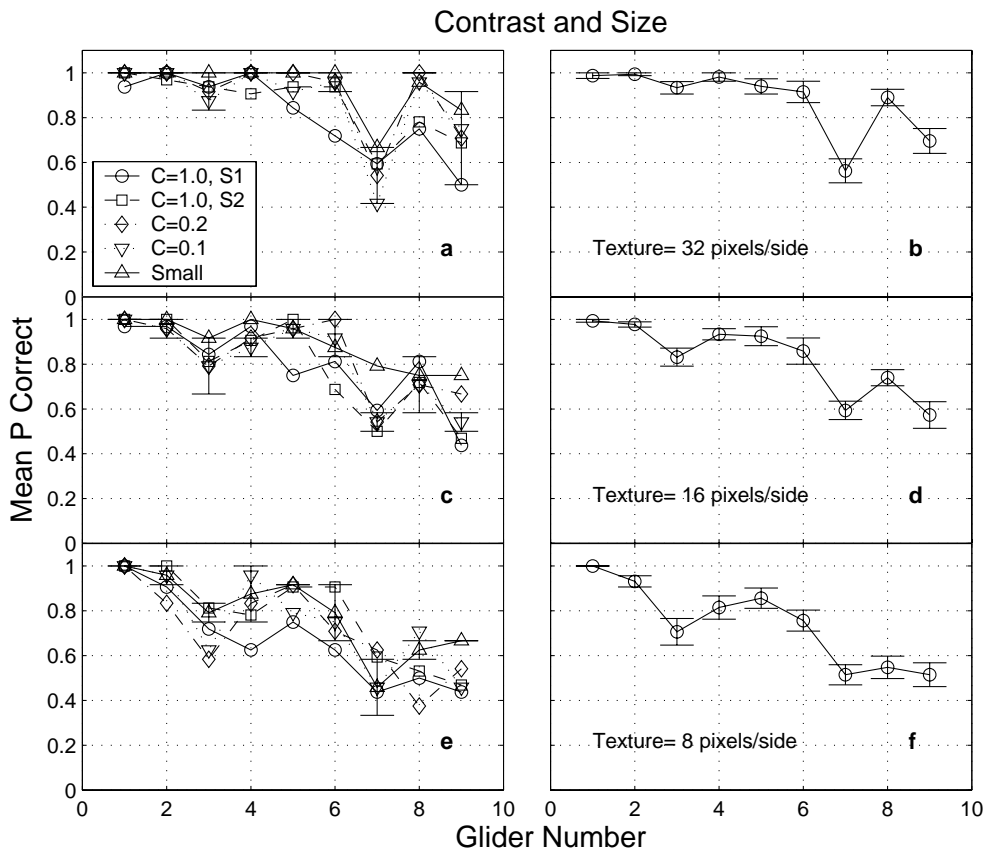


Figure 7

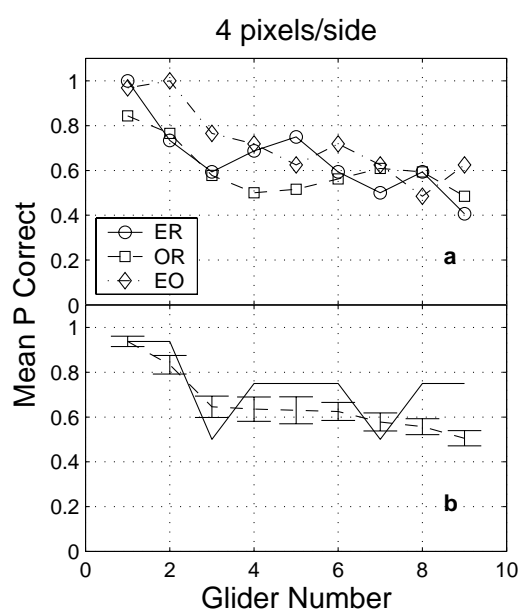
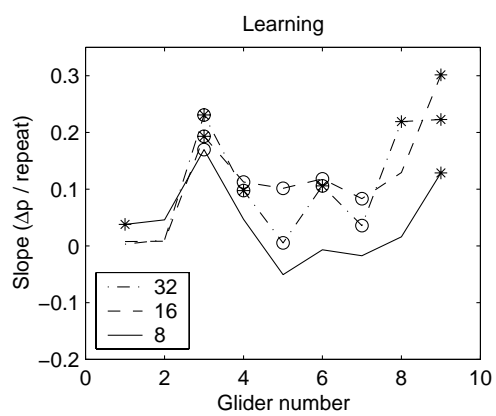


Figure 8

**Figure 9**

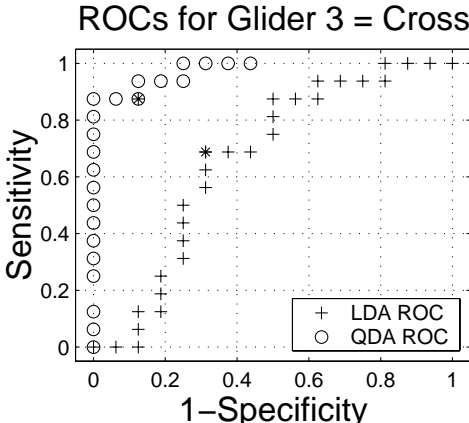


Figure 10

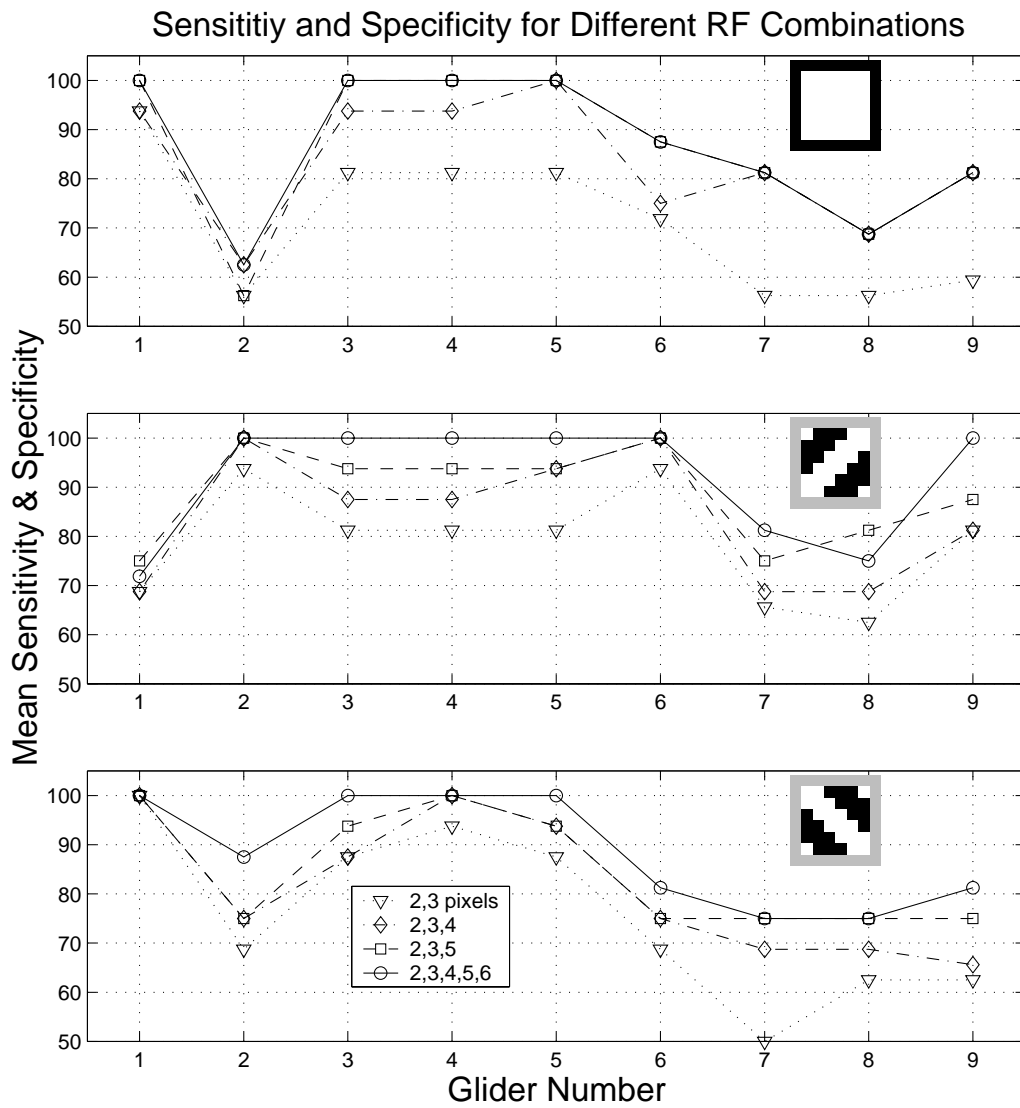


Figure 11

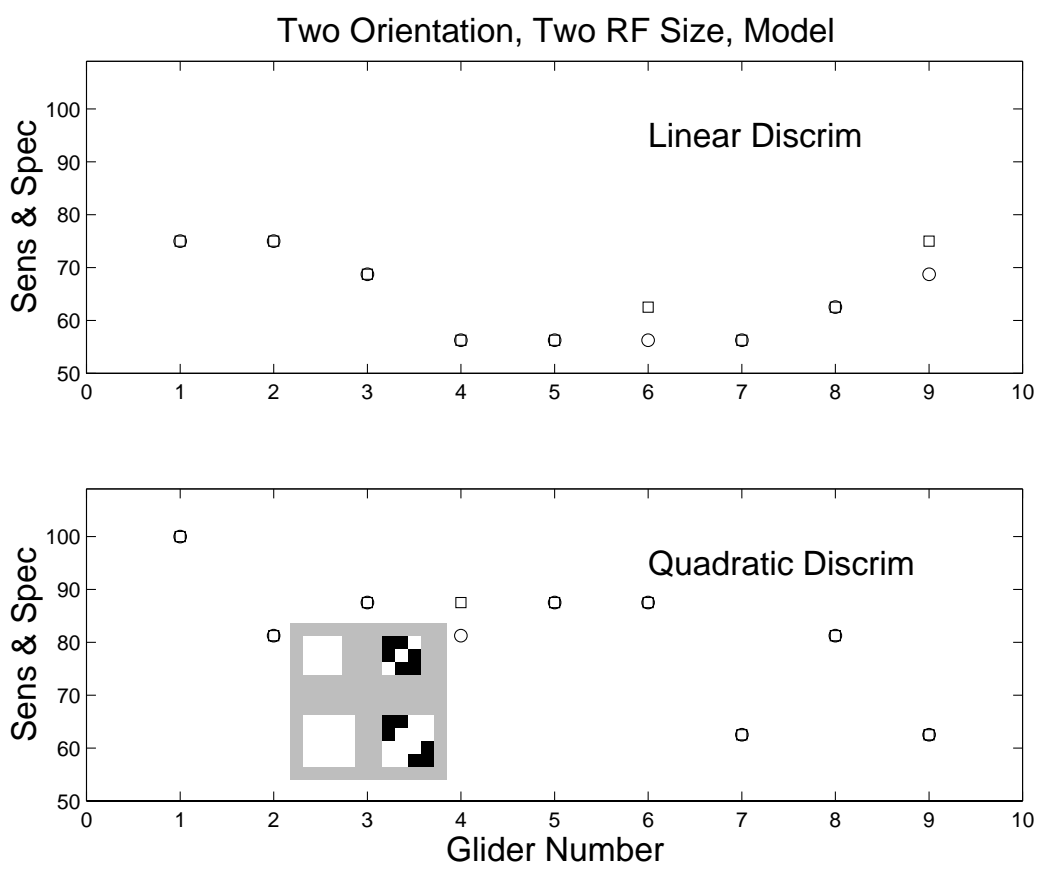


Figure 12

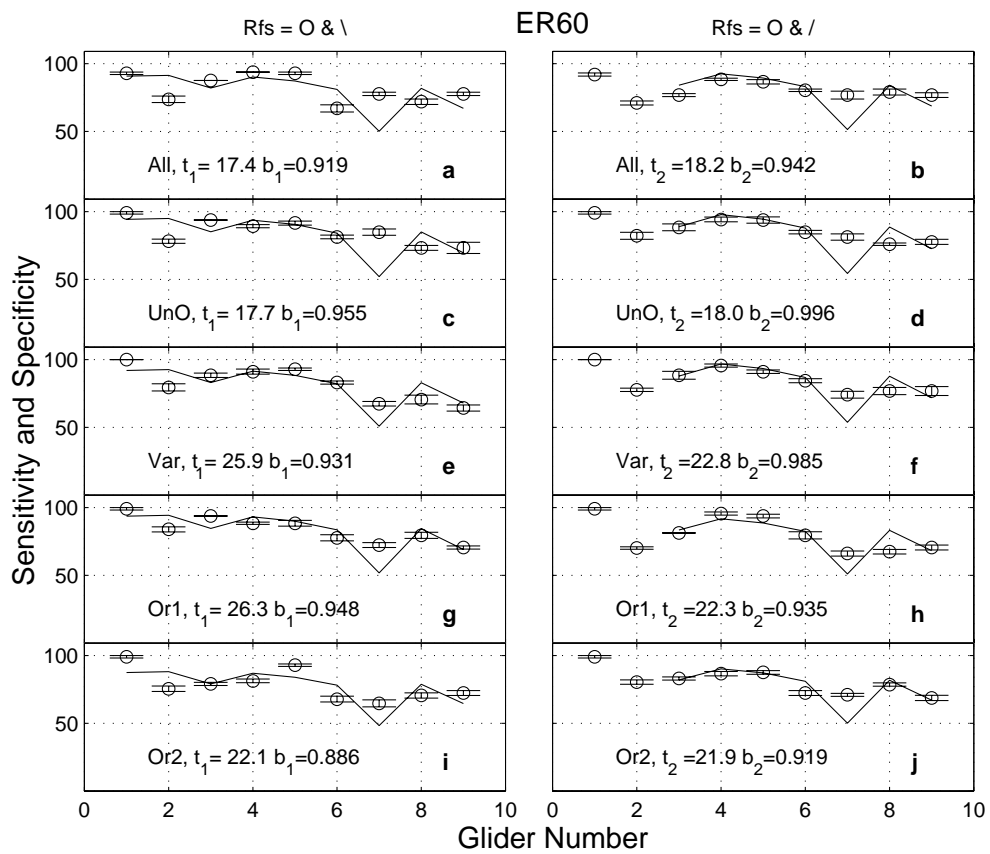


Figure 13

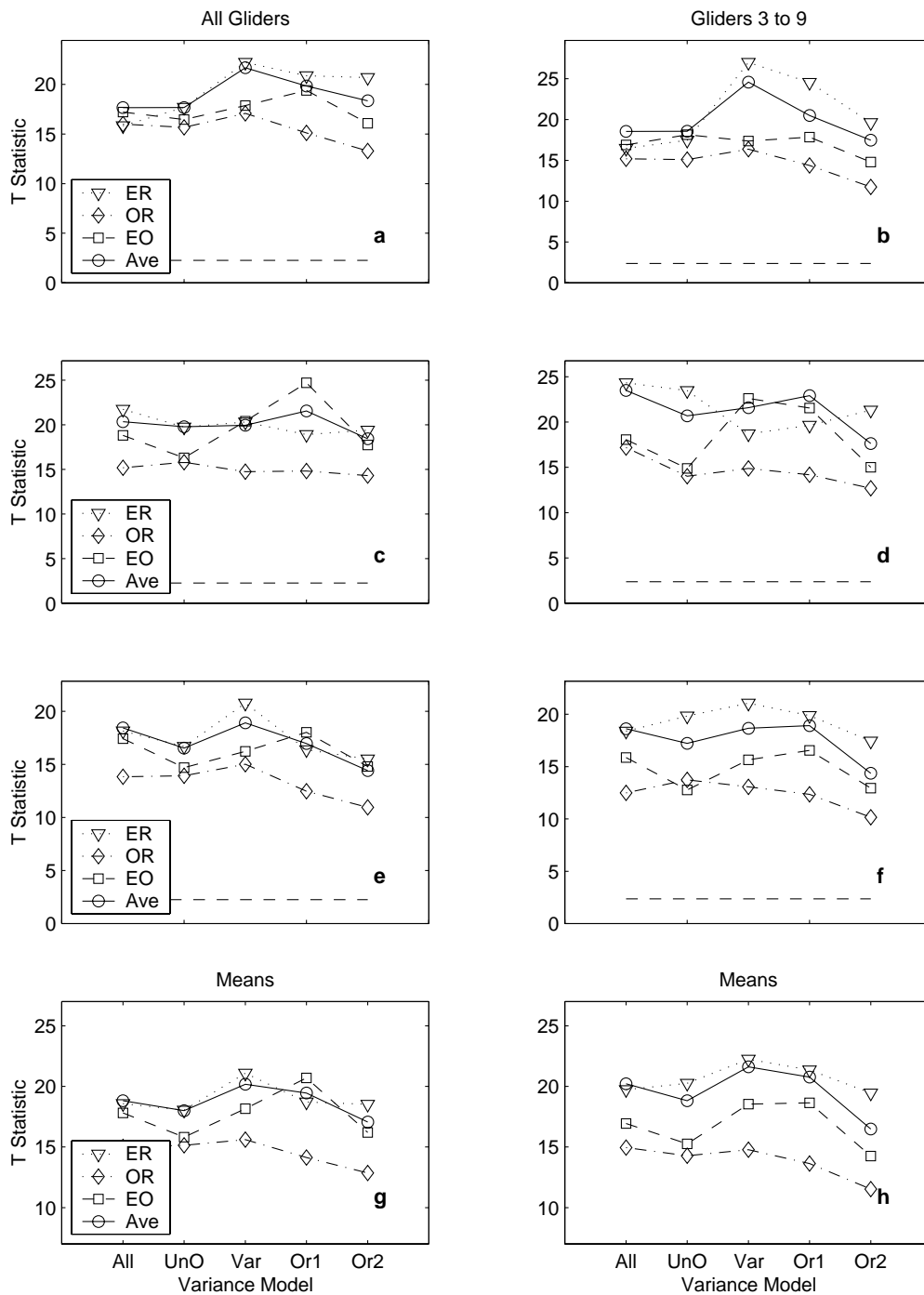


Figure 14

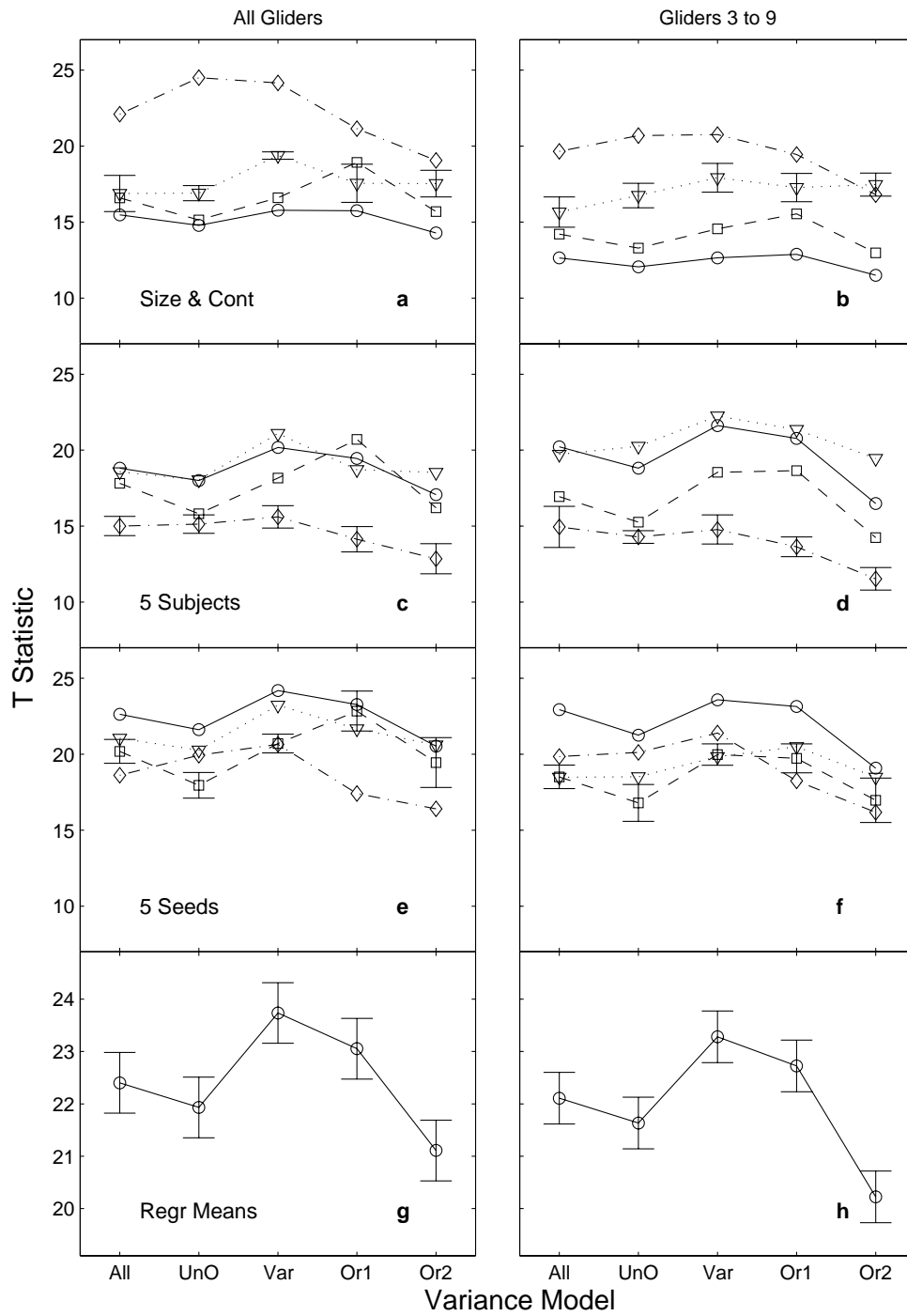


Figure 15

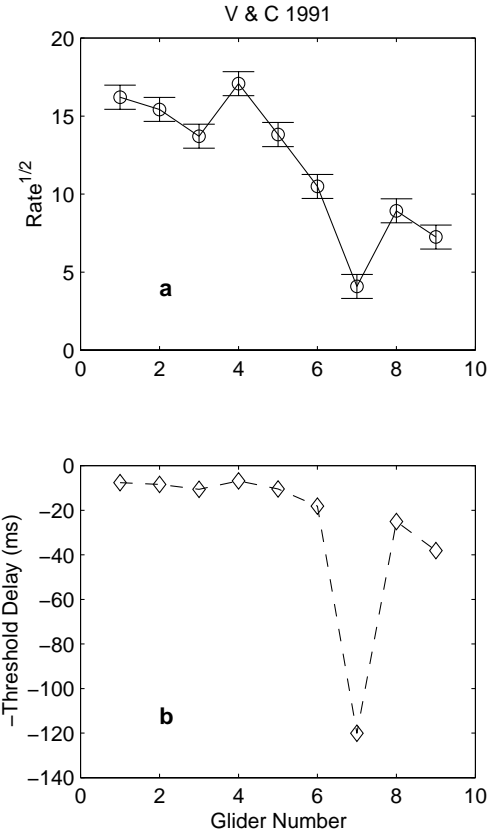


Figure 16

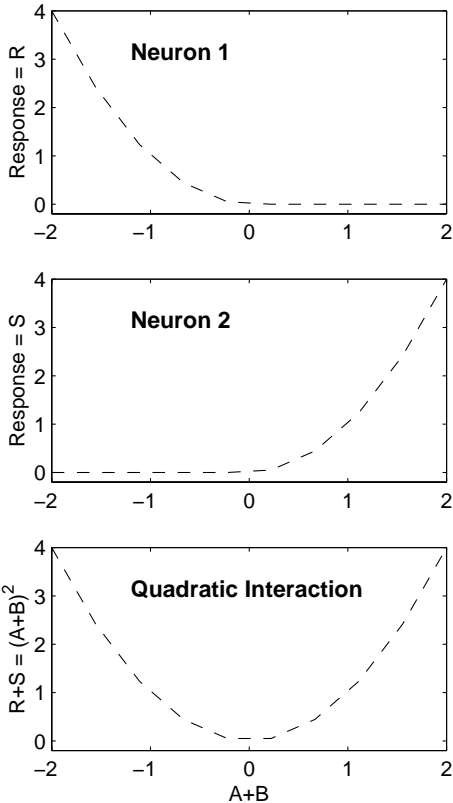


Figure 17

Appendix

If we have a NxM random binary texture R(N,M) having an equal probability of its two states, say $\{-1,1\}$ or $\{0,1\}$, then the possible number of different NxM textures $N(R(N,M)) = 2^{NM}$. In the case of the Box isotrigon textures (Odd or Even) the application of the Box glider, to a matrix of random $\{-1, 1\}$ elements means that the resultant textures are deterministic because in effect the state of three glider pixels determines the state of the fourth (Methods). For the Box glider we can summarise this by saying that the state $S(i+1,j+1)$ is a function of the three other pixels determined by the glider

$$S(i+1,j+1) = F(S(i,j), S(i+1,j), S(i,j+1))$$

In this case the resultant deterministic texture D(N,M) will have a number of possible states $N(D(N,M)) = 2^{N+M-1}$. This can be understood as follows. Assign the pixels of the first row and column randomly. We refer to these as *random pixels*. The other pixels are all assigned by the recursion rule determined by the Box glider. These are thus described as *deterministic pixels*. For example if we represent the Box glider as

$$\begin{matrix} * & * & n \\ * & d & n \\ n & n & n \end{matrix}$$

where the pixel d is determined by the three * pixel values. The null pixels, n, contribute nothing to d. Then we obtain a texture that can be shown as

	1	2	3	4	n
1	#	#	#	#	#
2	#	x	x	x	x
3	#	x	x	x	x
:	:	:	:	:		:
:	:	:	:	:		:
m	#	x	x	x	x

where #’s denote random boundary pixels and x’s denote deterministic pixels determined by the glider. Since the random pixels (#) take two values such as $\{-1, 1\}$ or

$\{0,1\}$, while the deterministic pixels are uniquely determined by the boundary random pixels, *then all possible generated textures are completely determined by the random boundary pixels*. Different textures are obtained by a different initial random pixel assignments. The number of random pixels (in the first row and column) is $n+m-1$ because the top left corner (1,1) is a degenerate pixel for counting. Hence, we know that the possible number of $N \times M$ Even/Odd textures for the Box glider is 2^{n+m-1} . Notice that the value 2 comes from the fact that each random pixel takes one of two possible values, and the state of any random pixel is determined independently to every other random pixel.

{Table A1 about here please}

For other glider textures, we count the possible number of textures in the same way except that we often have to assign more random pixels than just the first row and column. Notice that the same rules apply to every mini-texture of any larger parent texture, hence 3×3 mini-textures produced by the standard glider will have $2^{3+3-1} = 32$ possible states. As far as the standard Box textures are concerned this means that as N and M grow larger the possible textures represent a decreasingly small proportion of the possible random binary textures since $N(D(N,M)) / N(R(N,M)) = 2^{N+M-1} / 2^{NM} = 1 / 2^{(N-1)(M-1)}$. For other gliders we can ascertain the $N(D(N,M))$ by inspection of the glider rule and counting the minimum number of necessary random pixels. The glider rules and the resultant number of deterministic textures for a $N \times M$ pixel texture are summarised in Table A1.