PROCEEDINGS OF THE ROYAL SOCIETY

B | BIOLOGICAL SCIENCES

Cross- and auto-correlation in early vision

Horace Barlow and David L. Berry

References	**This article cites 26 articles, 6 of which can be accessed free** http://rspb.royalsocietypublishing.org/content/early/2010/12/04/rspb.2010.2170.full.html#ref-list-1
P<P	Published online 8 December 2010 in advance of the print journal.
Subject collections	Articles on similar topics can be found in the following collections neuroscience (437 articles) behaviour (1762 articles) cognition (436 articles)
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Proc. R. Soc. B* go to: **http://rspb.royalsocietypublishing.org/subscriptions**

Cross- and auto-correlation in early vision

Horace Barlow[1],* and David L. Berry[2]

[1]*Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3EG, UK*
[2]*Departamento de Física, Universidade de Évora, 7002 554 Évora, Portugal*

Neurons that respond selectively to the orientation of visual stimuli were discovered in V1 more than 50 years ago, but it is still not fully understood how or why this is brought about. We report experiments planned to show whether human observers use cross-correlation or auto-correlation to detect oriented streaks in arrays of randomly positioned dots, expecting that this would help us to understand what David Marr called the 'computational goal' of V1. The streaks were generated by two different methods: either by sinusoidal spatial modulation of the local mean dot density, or by introducing coherent pairs of dots to create moiré patterns, as Leon Glass did. A wide range of dot numbers was used in the randomly positioned arrays, because dot density affects cross- and auto-correlation differently, enabling us to infer which method was used. This difference stems from the fact that the cross-correlation task is limited by random fluctuations in the local mean density of individual dots in the noisy array, whereas the auto-correlation task is limited by fluctuations in the numbers of randomly occurring spurious pairs having the same separation and orientation as the deliberately introduced coherent pairs. After developing a new method using graded dot luminances, we were able to extend the range of dot densities that could be used by a large factor, and convincing results were obtained indicating that the streaks generated by amplitude modulation were discriminated by cross-correlation, while those generated as moiré patterns were discriminated by auto-correlation. Though our current results only apply to orientation selectivity, it is important to know that early vision can do more than simple filtering, for evaluating auto-correlations opens the way to more interesting possibilities, such as the detection of symmetries and suspicious coincidences.

Keywords: auto-correlation; symmetry; pinwheels; simple cells; complex cells; cortical hierarchy

1. INTRODUCTION

In this paper, we report results from psychophysical experiments using human observers that strongly suggest there are two different mechanisms for discriminating orientation in early vision. Hubel & Wiesel [1,2] discovered that most of the neurons of V1 were highly selective, not only for the position in the visual field of a visual stimulus, but also for other properties: a particular neuron only responds to a visual stimulus whose orientation, direction of motion and spatial frequency content lie within particular ranges whose mean values and widths vary from neuron to neuron. As well as discovering the orientation and motion selectivity of neurons in primary visual cortex, Hubel & Wiesel [1,2] also noted that there were two classes, which they named simple and complex cells. Subsequent work on different species, and work using awake, behaving, animals, has complicated the picture, but it is generally agreed that simple cells have smaller receptive fields than complex cells, have more direct inputs from LGN afferents and behave in a linear manner that can be explained in terms of the excitatory and inhibitory zones of their receptive fields [3,4]. It is also widely accepted that the simple cells act as tuned spatial filters having a wide range of peak sensitivities to position, orientation and spatial frequency, and that they do this by having neurons with a Gabor-like receptive field that are cross-correlated with the patches of the image that overlie each receptive field.

Complex cells tend to have larger receptive fields and to combine information from different parts of them in more complex ways [5–8]; they are more often directionally selective to motion than simple cells, and may be divisible into many different subcategories, possibly better described as a range of hybrids rather than just two contrasting types ([9]; but [10] do not fully subscribe to this interpretation). Hubel & Wiesel introduced the idea that the complexity of receptive field properties was increased in V1 and other visual areas by repeated transformations of the type exemplified by the simple cell/complex cell transition [2,7], though they recognized that the inability to be more precise about the nature of such hierarchical transformations made it seriously incomplete. We think this is the crux of the problem, and it has added interest because the discovery of pinwheels [11,12] is leading to a wealth of new information about the connections between neurons in V1; these organized connections must surely be doing more than the very simple task allotted to them in the diagrams illustrating Hubel & Wiesel's hierarchy, but what does this extra work achieve?

A vast amount of information about the physiology and anatomy of V1 has now been acquired, and has been incorporated in neural models of early vision, such as energy models of motion analysis [13–15] and normalization models of cortical responses [6]. These are satisfactory in some ways, but not in others. We cannot embark on a fair, critical, review, but our approach is largely motivated by the failure of the above models to provide insight into early vision at the level of computational theory. David Marr [16] suggested this was the

* Author for correspondence (hbb10@cam.ac.uk).

first, and clearly in his view the most important, of the three levels at which information-processing mechanisms have to be understood; he defined it as the level asking 'What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?'.

We thought it might be possible to answer some of these questions by psychophysical experiments on human subjects, and the first question we chose to ask concerned the mechanism of selectivity for orientation. We think this is the most distinctive of the new types of pattern selectivity to appear in V1, and therefore designed psychophysical tests that we hoped would answer the following specific computational question: 'Does V1 use cross-correlation, or auto-correlation, to detect oriented streaks in a noisy image?' The first possibility arises from the widely accepted suggestion, described above, that simple cells act as oriented spatial filters. For auto-correlation the situation is different, because there is no well-tested physiological example of its occurrence. Reichardt [17] introduced the radical idea that it is an important mechanism for detecting motion in the early stages of the visual pathway in the beetle *Chlorophanus*, but later developments of this model have tended to regard it as cross-correlation in the spatio-temporal domain, thus obscuring the original connection with auto-correlation [13–15]. Glass [18] suggested that the moiré effects generated by coherently oriented pairs in otherwise randomly positioned dots were a direct demonstration that the human visual system also computes local auto-correlations at an early level; but for many, the absence of a well-tested physiological example remains a problem, even though this cannot be taken as proof that it does not occur.

The results of our psychophysical measurements turned out to be even more interesting than we expected, for they clearly indicate that either cross- or auto-correlation can be used, depending on how the appearance of orientation had been generated.

The distinction between cross-correlation and auto-correlation is crucial for our argument, so we repeat here the definitions from Haralick & Shapiro [19], and we shall follow this usage ourselves. Cross-correlation γ, normalized for two-dimensional images, is given by

$$\gamma = \frac{1}{N} \sum_{x,y} \frac{(I(x,y) - \bar{I})(T(x,y) - \bar{T})}{\sigma_I \sigma_T}, \qquad (1.1)$$

and it gives a measure of the similarity between a fixed, predetermined template $T(x,y)$ and the pixel values in a patch of the image $I(x,y)$ covering the same positions, each containing N pixels; $\bar{I}$ and $\bar{T}$ denote the means and σ_I and σ_T denote the standard deviations of $I(x,y)$ and $T(x,y)$, respectively. This is profoundly different from auto-correlation α, given by

$$\alpha = \frac{1}{N} \sum_{x,y} \frac{(I(x,y) - \bar{I})(I(x-u, y-v) - \bar{I})}{\sigma_I^2}. \qquad (1.2)$$

This gives a measure of the similarity between the pixel values, whatever they happen to be, of one patch of the image and those of another patch shifted by u in x and v in y. Notice that auto-correlation does not specify what it is that is similar. It can thus be considered as a

coincidence or symmetry detector, something that detects the abstract property of similarity or sameness, but does not discriminate between different examples of it: for these particulars, the response is invariant. These characteristics cannot be detected by cross-correlation alone, for it will be seen from expression (1.1) that the contributions to γ from every particular location in a pattern are influenced by the value at the corresponding position in the template, but are not influenced by the concurrent contributions from all other locations in the pattern. Two input patterns that differed only in the joint input values at two different locations would necessarily have the same total value of γ, for none of the separate contributions to γ are determined by joint values at pairs of input positions.

The difference between these two operations (sometimes called first and second order for cross- and auto-correlation, respectively) is important for understanding why changes in the mean dot density of random dot patterns result in such different effects. It is because, for cross-correlation, only one of the variables whose products are summed has a random component, whereas for auto-correlation, both have. Thinking along these lines has its antecedents in Bela Julesz's conjecture that humans cannot distinguish between textures with identical second-order statistics, which was made in 1962. He later claimed that he had disproved it [20], but we think it set the fashion for considering the order of the statistical analysis required for perceptual discriminations, and that it is correct in the slightly modified form we shall give later.

(a) Choosing the tasks

The plan of our experiment was to measure the ability of the human visual system to discriminate regular, patterned features presented in arrays of randomly positioned dots. The threshold strength of the signal for the regular feature changes when the mean density of the randomly positioned dots in the array is changed, but the expected relationship between the two differs according to whether cross- or auto-correlation is used for the discrimination. Conversely, we can infer which mechanism is being employed by observing which relationship is being followed. We therefore wanted two test tasks having the following appearances and ranges of adjustment:

(a) they should look as similar as possible, for if they were readily distinguishable by subjective appearance at threshold stimulus strengths, the change of method from cross- to auto-correlation or vice versa might involve higher cognitive processes;

(b) the signal strength of the test patterns must be adjustable in order to determine thresholds;

(c) the noise that limits detection by cross- or auto-correlation must also be variable; for both tasks, we achieved this by varying the initial mean numbers of dots allotted to each pixel position in the arrays, before introducing the regular feature whose threshold was to be measured.

We chose to test cross-correlation by measuring thresholds for detecting the oriented streaks induced by low-amplitude sinusoidal spatial modulations of dot density in random dot patterns, partly because we already knew that this followed the inverse square-root of unmodulated dot density over a considerable range, as expected

theoretically for cross-correlation (see eqn (8) in [21]), but also because we suspected that the streakiness would look very similar to that induced by coherent dot pairs in Glass patterns [18]. These moiré patterns would form a natural partner-task for testing auto-correlation, and it has already been found that, over a limited range, the coherence threshold stays nearly constant with changes of dot density, as expected if auto-correlation is used [22,23].

Some of the test patterns generated for these tasks are shown in figure 1 and illustrate that, at signal strengths just above threshold; it is not easy to discriminate between streaks caused by dot density modulation and streaks caused by Glass pairs. It is also clear that there are no problems with grading signal strength; therefore, requirement (a) and (b) above are at least partially met. For the noise level (c), we have assumed that the standard deviation of dot density, when this is expressed as the mean number of dots per pixel, λ, is equal to the square root of this mean, and this was confirmed by our simulations for all the conditions encountered in these experiments (see §2). Note, however, that the noise when discriminating by auto-correlation is not the same as the noise when discriminating by cross-correlation, because it depends upon the standard deviation of the number of spurious pairs having properties indistinguishable from the deliberately introduced coherent pairs, rather than upon the standard deviation of local dot density.

When calculating theoretical limits and doing computer simulations, we are primarily interested in how well a proposed mechanism or model collects together and uses the statistical information that the visual system has been presented with and is needed to perform the task. What matters, then, is not simply how signal strength at the output varies with signal strength at the input under different conditions, but how signal-to-noise ratios attainable at the output compare with the signal-to-noise ratios attainable at the input under different conditions. It is the statistical efficiency of the process that matters, not just the signal strength. We have, however, kept our simulations as simple as possible, and have not introduced factors such as the intrinsic noise of neurons; we do not doubt that there are conditions where this is important, but for simplicity we chose in the first place to avoid such conditions.

The usual method of generating noisy dot arrays using randomly positioned dots has a serious limitation that we call occlusion: except at very low dot densities, a significant fraction of the dots programmed for a particular location will have no effect, because that location has already been occupied by a previously programmed dot. The way this was overcome by using graded dot intensities will be described in §2.

2. METHODS

(a) Observers and apparatus

The procedures used were entirely non-invasive and were approved by the local ethics committee. The observers were the two authors plus another experienced psychophysicist, all having normal or corrected-to-normal vision. Aided by a chin rest, observers viewed the stimuli binocularly at 100 cm in a darkened, quiet, room. The stimuli were circular with a diameter of 200 pixels and were presented on a 19 inch CRT monitor (CTX EX951F) (1024 × 768 pixels,
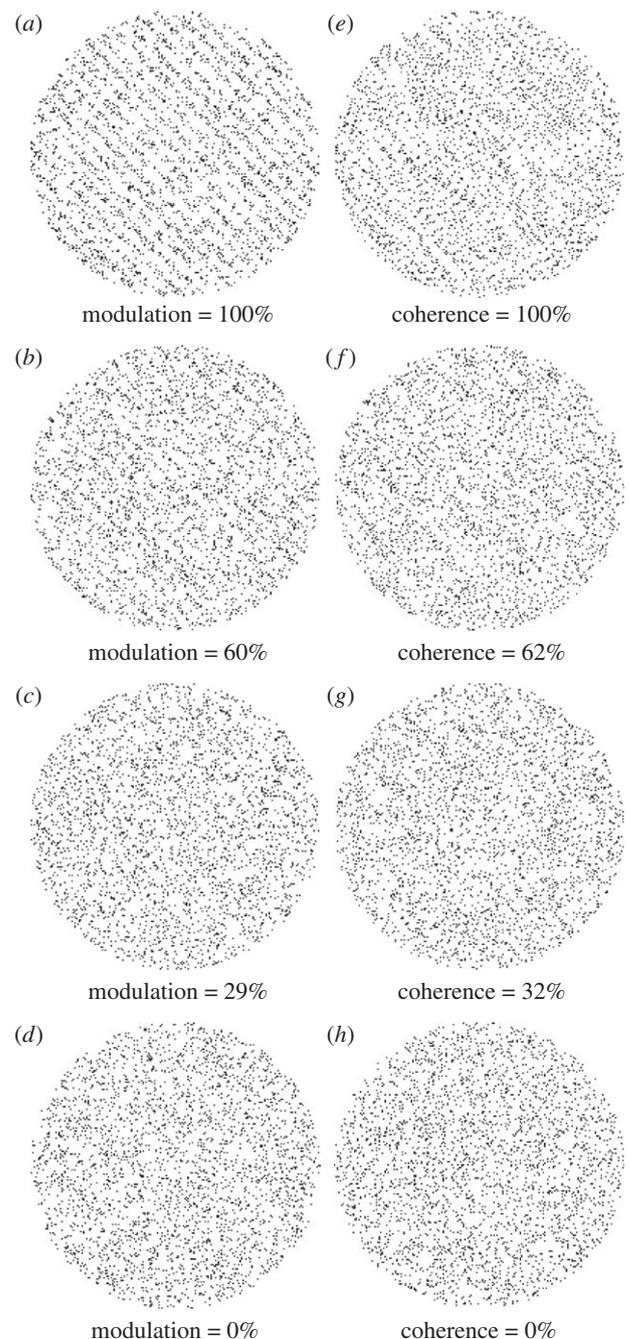
Figure 1. The similarity of streaks produced by weakly modulated gratings and low coherence Glass patterns. The gratings (a–d), produced by sinusoidally modulating the expected number of dots per pixel position in arrays of randomly positioned dots, and the Glass patterns (e–h), produced by coherently pairing a fraction of dots in arrays of otherwise randomly positioned dots, show easily distinguishable streaks at high values of modulation and coherence, but these become practically indistinguishable at values approaching threshold (19.75% modulation for gratings, 24.1% coherence for Glass patterns). To improve reproduction, black dots on a white background are used instead of the white or graded dots on a black background used in the experiments. A mean value, λ, of 0.1 dots per pixel was used with the streaks tilted 45° left of vertical. The patterns should be observed at ca 50 cm to give the same spatial frequency and dot separation as those in the experiments (see §2). Any streaks observed in the patterns at zero modulation or coherence are due to unintended dot alignments occurring by chance, and these set the theoretical lower limit for streak detection.

0.34 mm [H] × 0.34 mm [V] per pixel, 85 Hz frame rate). The maximum luminance of the monitor was 92.4 cd m^{-2} and was calibrated and linearized by constructing a look-up table. The programs were written in MATLAB using the psychophysics toolbox extension [24].

(b) *Stimulus generation*

Examples of the types of patterns used in the experiments are shown in figure 1. The dots making up these patterns occupy a single pixel position within a circular region of 200 pixels diameter and had luminances controlled by an 8 bit D/A converter. Dot densities are expressed as the mean expected number of dots per pixel position. To generate the gratings, each pixel within the array was designated a probability determined by the eventual two-dimensional sinusoidal pattern required, and numbers of dots determined for the required values of mean dot density were then distributed accordingly within the array. These gratings, with a spatial frequency of 7.44 cycles per degree, were oriented at 45° either to the left or right of vertical. To generate the Glass patterns, pairs of dots, with a dot separation of 8.3 arcmin, were placed at random within the circular region, with the pairs tilted at 45° either to the left or right of vertical.

In the initial experiments using only white dots, each pixel in the array held either zero or one dot, corresponding to black or white. For the main psychophysical experiments and simulations, we used pixels whose luminances could be varied, having values increasing linearly with the number of times a particular pixel position had been chosen to contain a dot, and these we called 'graded dots'. For each value of mean dot density, M_p, we calculated the expected standard deviations of the number of dots per pixel, σ_p, and devoted the whole range of pixel luminances to cover the range ($M_p \pm 2 \times \sigma_p$). The output of the D/A converter was set to produce a luminance increasing linearly with the number of dots held by each particular pixel. Thus, the pixels occupied by ($M_p + 2 \times \sigma_p$) or more dots were set at white, those occupied by ($M_p - 2 \times \sigma_p$) or less dots were set to black and intermediate values were assigned the appropriate grey value. This assignment of dot luminances was chosen in order to improve the fidelity of the visual system's computations of cross- and auto-correlations, on the assumption that the system responds linearly to luminances in the range covered by the graded dots.

(c) *Threshold determination*

We used a modified staircase procedure. Up to eight staircases for eight different mean dot densities were run simultaneously in random order until 20 stimuli had been presented on each staircase. A small fixation dot positioned at the centre of the stimulus area, together with a short beep, occurred 500 ms before each stimulus, which lasted 160 ms. The observer indicated on the keyboard whether he or she saw left or right tilting streaks, or was uncertain, and received auditory feedback for correct or incorrect responses, but none for signals of uncertainty. The response time was unlimited, but in practice lasted a second or two. Reversal points on each staircase were corrected by half a step value according to the direction of the reversal, and these were used to estimate thresholds and their standard errors. There were usually 6–12 reversals on a staircase, and if the false-positive rate was higher than 5 per cent, or the number of reversals on a staircase was below 6, the results were rejected. Other details of the procedure were conventional.

(d) *Simulations*

In the simulations presented here, we generated the dot arrays in the same way as they had been generated for the gratings and Glass patterns in the experiments. We then set the parameters for cross- and auto-correlation at their optimum values for the tasks in hand. For cross-correlation, this required the template to be matched exactly to the size, shape, spatial frequency, phase and orientation of the test patterns, and for auto-correlation, it required the complete test pattern to be multiplied by its copy shifted by one Glass pair separation in the direction of the separation. The simulation programs computed decision variables at narrowly spaced intervals of dot density for the expressions γ and α in equations (1.1) and (1.2) for samples of gratings and Glass patterns. For comparison with experimental thresholds, we determined the modulation or coherence for which the discriminability index d' given by $d' = (M_{SN} - M_N)/\sigma_N$, is equal to 2, where M_{SN} is the mean of the distribution of values of cross- or auto-correlations determined for sample populations at specified values of modulation or coherence. M_N and σ_N are the means and standard deviations of equivalent distributions of sample populations for zero modulation or coherence (i.e. noise-alone stimuli).

When the dot density is low and ungraded white dots are being used, it can be assumed that the numbers of dots in single pixels are Poisson distributed, and therefore the standard deviation of dot numbers in any single pixel, or designated subset of pixels, is equal to the square root of the mean dot number, or sum of dot numbers. However, some of the test arrays are large and contain significant fractions of all the dots, and they may not be simple sums of the dots in individual pixels, but weighted sums, which are not in general Poisson distributed. We have done many of the simulations in two ways, first using the assumption that the standard deviations of the noise-alone means are equal to the square roots of the mean numbers of dots they contain, and second by generating large samples of such distributions and measuring their standard deviations directly. The assumption that they would be equal to the square root of the mean number of dots turned out to be a good approximation, provided that only low values (below about 50%) of modulation and coherence were used. Thus, the assumption that the standard deviations of dot numbers in any test array is equal to the square root of the mean of the number of dots it contains is a good approximation *under the conditions of these experiments*.

3. RESULTS

Figure 2 shows the result of the first experiment: log thresholds were measured for amplitude-modulated gratings and varying coherence Glass patterns as functions of log(λ), using white, ungraded, dots. As generally expected from previous results (e.g. [21,23]), the results show slopes (using log/log coordinates) quite close to -0.5 for modulation thresholds at low dot densities, and quite close to zero for coherence thresholds also at low dot densities, but they both then deviate upwards from the expected values (see best-fitting lines of slopes -0.5 or 0 calculated for the first four points). The range of validity was thus limited to a 30-fold range of dot densities. At low values, it was subjectively clear to the observer that cognitive factors tended to be used for making the judgments, and this discouraged attempts to extend the test range downwards. However, the upper
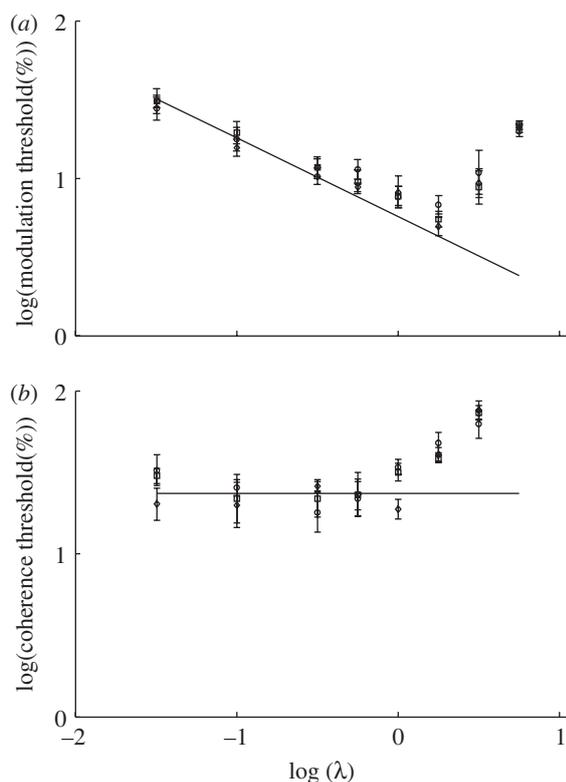
Figure 2. Log thresholds for detecting the orientations of streaks produced by (*a*) gratings or (*b*) Glass patterns, as functions of log dot densities, using ungraded, white, dots. The points are the thresholds for three subjects determined by a staircase method, with error bars equal to $2\times$ s.e. The straight lines were given the theoretically expected slopes of (*a*) -0.5 and (*b*) 0, and were computed best fits to the first four points only. They clearly become unsatisfactory fits when $\log(\lambda)$ reaches about -0.5, thus limiting to about 30-fold the range of dot densities over which the theoretical expectations are followed. See text for further discussion.

limit for the range of validity appears to be due to a quite different effect—that of occlusion, explained below.

In figure 2, the upward deviations from the predicted relations start when $\log(\lambda) \approx -1$, and are easily signifi-cant when $\log(\lambda) \approx 0$, corresponding to values for λ itself of 0.1 and 1, respectively. With our method of gen-erating white dots, if two or more dots are allocated to a pixel, still only one is shown, and any excess allocations are occluded, having no effect whatever on the display. Calculation shows that the probability of two or more dots occurring in a single pixel exceeds 0.05 when $\lambda \geq 0.355$, or $\log(\lambda) \geq -0.45$, and since occluded dots cannot help the observer to see oriented streaks, occlusion provides a satisfactory explanation for the experimental thresholds rising at this point.

To test this explanation, figure 3*a*,*b* show threshold measurements like those of figure 2, but using graded dots. The results using ungraded white dots have been copied over from figure 2. Two points are immediately obvious: first, up to a value of $\log(\lambda) \approx -0.5$, the results using graded dots agree well with those using white dots; second, the results using graded dots demonstrate good approximations to the predicted inverse square root and constant coherence threshold relationships far above the value of $\log(\lambda) \approx -0.5$, the agreements both holding over 1000-fold ranges of dot density.
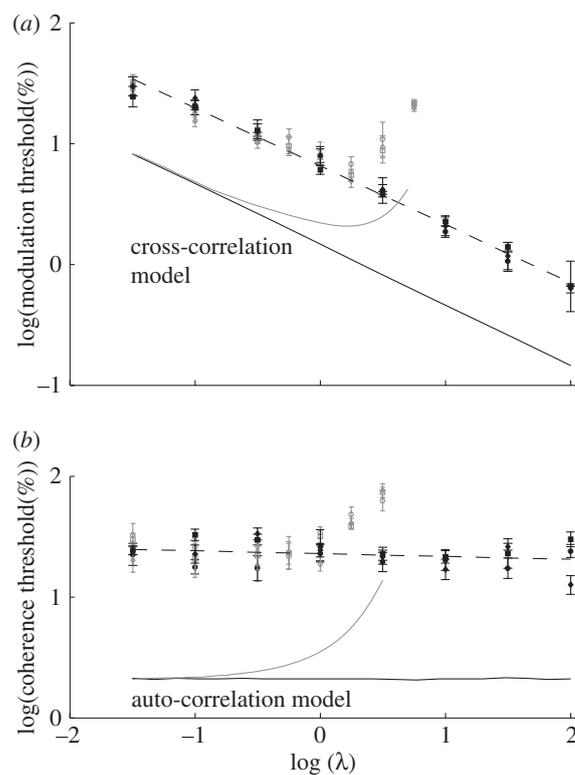

Figure 3. Same as figure 2, but using graded dots. When using graded dots, log threshold versus log dot density relations for (*a*) gratings and (*b*) Glass patterns become straight lines of characteristic slope (dashed regression lines and filled symbols) that match those of ideal models using cross- and auto-correlation (black lines). Threshold results for ungraded white dots are carried over from figure 2 (unfilled symbols), and results for ideal models using ungraded white dots (grey lines curving upwards) are also shown. The means and standard errors of the slopes of the main regression lines are -0.482 ± 0.016 and -0.024 ± 0.010, close to those expected theoretically (-0.5 and 0). The ideal models using graded dots accurately follow the theoretical slopes of -0.5 and 0. For ideal detectors, the area of the detector was matched to the area of the target stimulus, but no such matching was possible for the psycho-physical results; more information is therefore needed to estimate absolute efficiencies (see text).

Figure 3 also shows the results of simulations for ideal detectors using cross-correlation and auto-correlation (black lines). For graded dots, these have slopes of -0.5 and 0, accurately following the theoretical predictions. The grey lines curving upwards show the results for ungraded white dots, confirming that the simu-lations, like human observers, are affected by occlusion when the targets use ungraded white dots.

We think the slopes for ideal cross-correlation and auto-correlation in figure 3 add support to the hypothesis that either method can be used for detecting the orien-tations of streaks. The fact that avoiding occlusion has such a dramatic effect in extending the ranges for which these relationships hold, and the fact that it works for the ideal simulations as well as for the human observers, strengthens the case.

As shown in figure 3, our current simulations have thresholds about 25 per cent of those of our observers for cross-correlation, and about 8 per cent of those for auto-correlation. This could be taken to indicate that

our ideal simulations are able to do a much better job of interpreting the information in the stimulus dots than the real visual system is able to do. This would, however, be misleading, because the ideal systems have an advantage over the natural system: their parameters are matched to the parameters of the stimulus, whereas the real system's parameters cannot be adjusted by the experimenter. An alternative would be to match the parameters of the stimulus to those of the natural system but we do not yet know their values. The main such parameter is likely to be the target area over which the information contained in the positions of the dots can be combined efficiently to give the correct impressions of the orientations of the streaks. This area is likely to be much less than the 15 deg^2 provided by the full test area, so when making the comparison of figure 3, the ideal system has a great advantage over the real system because it uses the whole test area. This advantage should be reduced when a smaller test area is used, and this was found to be the case, but further experiments are needed to obtain proper estimates of optimal statistical efficiencies and the conditions under which they are obtained.

4. DISCUSSION
(a) *Comparisons with earlier results*
The psychophysical results of figure 3 clearly indicate that early vision uses cross-correlation to detect the oriented streaks produced in random dot arrays by sinusoidal spatial modulation of dot density, whereas it uses auto-correlation to detect the oriented streaks produced by coherent pairs, as in translational Glass patterns. The earlier studies referred to above had suggested that this was the case, but the technique for using graded dots to avoid occlusion at high dot densities leads to much more convincing experimental results. Notice that the technique for generating the graded dot displays was designed to improve the fidelity of the visual system's computations of cross- and auto-correlations, so the fact that it works so well adds confidence to our belief that our psychophysical procedures genuinely test the use of these two types of computation.

We shall not understand the full significance of there being two computational methods for distinguishing orientation until we find out whether there are similar dichotomies for the other main computations thought to occur in early vision—those for motion, stereopsis and perhaps colour. It is also important to determine whether, as we believe, they both occur at a single level, or whether the two components occur at successive levels in early vision, for example V1 for cross-correlation and V2 for auto-correlation, but a conclusive answer to this question is likely to need a neuro-physiological approach.

As far as we know, there are no experimental results that conflict directly with ours, though some may at first appear to do so. Dakin [25], for example, suggests that cross-correlation alone can detect Glass patterns, but he does not appear to have made quantitative comparisons to support his claim. It is also not clear whether his tests were aimed (as ours were) at the first of the two stages involved in detecting Glass patterns—that of detecting coherent pairs—or at the second stage, that of detecting whether the pattern of coherence is translational, radial, circular or hyperbolic. There is also a lack

of positive evidence for auto-correlation in the neuro-physiological analysis of responses to paired dots by Smith *et al.* [26], but while these results show good evidence for cross-correlation, we do not think that any of them challenge our evidence that auto-correlations are also computed in early vision.

We are not aware of any other studies that have tested for the occurrence of auto-correlation in early vision by the method described here. It would certainly be interesting to know whether measurements of signal-to-noise ratios in energy models [13–15] can mimic the results shown in figure 3.

(b) *Significance of auto-correlation*
If auto-correlation, or an equivalent wavelet computation, does occur, this must be important for it makes possible the discrimination of classes of patterns defined by joint properties of pairs of pixels. As pointed out in §1, the products that contribute to the value of a cross-correlation are all between single pixel values and a pre-specified template-weighting factor: the contributions of a particular pixel do not vary according to the simultaneous contribution of other pixels, and cross-correlation alone is therefore blind to second-order statistical differences between patterns of pixels. On the other hand it is quite different for auto-correlations; they are formed solely from the products of the values of two image pixels, so differences in the frequencies of joint events are certainly expected to cause distinguishable differences in α.

This has an enormous effect in increasing the total number of patterns that could, in principle, be distinguished from each other by auto-correlation, compared with the number distinguishable by cross-correlation alone, and this immediately suggests an explanation for the explosive increase in the number of neurons per incoming afferent neuron that occurs when the visual pathway enters the cortex. This explosive increase has long been known to occur [27], and the ratio of cortical neurons to input afferent fibres has more recently been shown to reach the astonishing figure of $10\,000:1$ in the foveal projection of macaques [28].

Surely this explosive expansion must indicate that a radically different form of computation is being employed, and we think this is likely to be the systematic exploitation of auto-correlation in the cerebral cortex to discover second-order features in input patterns. In other words, the conjecture Julesz [20] made in 1962 and later withdrew was very nearly right: he should have said 'The general ability to distinguish between patterns with different second-order statistics only becomes possible in the cerebral cortex, through its systematic use of auto-correlation'.

Because of the expansion, it is reasonable to expect that there is a huge number of different complex cells each connected to the same, or nearly the same, set of simple cells, but differing in the detailed pattern of their connections. If these differences amount to specifying different displacements (u,v) in equation (1.2) for a large number of different auto-correlations α, then we would have an array of elements that is, in effect, looking for pre-specified auto-correlations in the filtered input from the image, and these might be the functional correlates of the geometrical arrangements that Ben-Shahar & Zucker [29] describe.

Our search for the computational goal of early vision has produced some evidence that auto- as well as cross-correlation is used in the pinwheels of V1 to detect higher order regularities in the visual input, such as symmetries and suspicious coincidences. Perhaps William James [30] was right in claiming that, in his words, 'the sense of sameness is the very keel and backbone of our thinking', and in that case the computational goal of V1 may turn out to be closer to that of the cerebral cortex as a whole than has been generally recognized.

REFERENCES

1 Hubel, D. H. & Wiesel, T. N. 1959 Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–591.

2 Hubel, D. H. & Wiesel, T. N. 1962 Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154.

3 Marcelja, S. 1980 Mathematical description of the response of simple cortical cells. *J. Opt. Soc. Am.* **70**, 1297–1300. (doi:10.1364/JOSA.70.001297)

4 Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. 1978 Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol.* **283**, 53–77.

5 Carandini, M. 2006 What simple and complex cells compute. *J. Physiol.* **577**, 463–466. (doi:10.1113/jphysiol.2006.118976)

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**, 8621–8644.

7 Hubel, D. H. 1982 Evolution of ideas on the primary visual cortex, 1955–1978: a biased historical account. *Nature* **299**, 515–524. (doi:10.1038/299515a0)

8 Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. 1978 Nonlinear spatial summation in the receptive fields of complex cells in the cat striate cortex. *J. Physiol.* **283**, 78–100.

9 Mechler, F. & Ringach, D. L. 2002 On the classification of simple and complex cells. *Vis. Res.* **42**, 1017–1033. (doi:10.1016/S0042-6989(02)00025-1)

10 Priebe, N. J., Mechler, F., Carandini, M. & Ferster, D. 2004 The contribution of spike threshold to the dichotomy of cortical simple and complex cells. *Nat. Neurosci.* **7**, 1113–1122. (doi:10.1038/nn1310)

11 Blasdel, G. G. & Salama, G. 1986 Voltage sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* **321**, 579–585. (doi:10.1038/321579a0)

12 Bonhoeffer, T. & Grinvald, A. 1991 Orientation columns in cat are organized in pin-wheel like patterns. *Nature* **353**, 429–431. (doi:10.1038/353429a0)

13 Adelsen, E. A. & Bergen, J. R. 1985 Spatio-temporal energy models for the perception of motion. *J. Opt. Soc. Am. A* **2**, 284–299. (doi:10.1364/JOSAA.2.000284)

14 van Santen, J. P. H. & Sperling, G. 1985 Elaborated Reichardt detectors. *J. Opt. Soc. Am. A.* **2**, 300–321. (doi:10.1364/JOSAA.2.000300)

15 Watson, A. B. & Ahumada, A. J. 1985 Model of human visual-motion processing. *J. Opt. Soc. Am. A.* **2**, 323–342. (doi:10.1364/JOSAA.2.000322)

16 Marr, D. 1982 *Vision.* San Francisco, CA: W H Freeman and Co.

17 Reichardt, W. 1961 Autocorrelation, a principle for the evaluation of sensory information. In *Sensory communication* (ed. W. A. Rosenblith), pp. 303–317. New York, NY: Wiley.

18 Glass, L. 1969 Moirée effect from random dots. *Nature* **223**, 578–580. (doi:10.1038/223578a0)

19 Haralick, R. M. & Shapiro, L. G. 1992 *Computer and robot vision,* vol. II, pp. 316–317. Reading, MA: Addison-Wesley.

20 Julesz, B. 1981 Textons, the elements of texture perception, and their interactions. *Nature* **290**, 91–97. (doi:10.1038/290091a0)

21 van Meeteren, A. & Barlow, H. B. 1981 The statistical efficiency for detecting sinusoidal modulation of average dot density in random figures. *Vis. Res.* **21**, 765–777. (doi:10.1016/0042-6989(81)90174-7)

22 Downing, C. & Movshon, J. A. 1989 Spatial and temporal summation in the detection of motion in stochastic random dot display. *Invest. Ophthalmol. Vis. Sci.*, **72**(Suppl. 30), 72–87.

23 Maloney, R. K., Mitchison, G. J. & Barlow, H. B. 1987 The limit to the detection of Glass patterns in the presence of noise. *J. Opt. Soc. Am. A* **4**, 2336–2341. (doi:10.1364/JOSAA.4.002336)

24 Brainard, D. H. 1997 The psychophysics toolbox. *Spatial Vis.* **10**, 433–436. (doi:10.1163/156856897X00357)

25 Dakin, S. C. 1997 The detection of structure in Glass patterns: psychophysics and computational models. *Vis. Res.* **37**, 2227–2246. (doi:10.1016/S0042-6989(97)00038-2)

26 Smith, M. A., Bair, W. & Movshon, J. A. 2002 Signals in macaque striate cortical neurons that support the perception of Glass patterns. *J. Neurosci.* **22**, 8334–8345.

27 Chow, K.-L., Blum, J. S. & Blum, R. A. 1950 Cell ratios in the thalamo-cortical visual pathway of *Macaca mulatta*. *J. Comp. Neurol.* **92**, 227–239. (doi:10.1002/cne.900920208)

28 Hawken, M. J. & Parker, A. J. 1991 Spatial receptive field organisation in monkey V1 and its relationship to the cone mosaic. In *Computational models of visual processing* (eds M. S. Landy & J. A. Movshon), pp. 83–93. Cambridge, MA: The MIT Press.

29 Ben-Shahar, O. & Zucker, S. 2004 Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. *Neural Comput.* **16**, 445–476. (doi:10.1162/089976604772744866)

30 James, W. 1890 *Principles of psychology*, ch. 12. New York, NY: Holt.