Detection of bilateral symmetry using spatial filters

S. C. DAKIN and R. J.WATT*

Department of Psychology, University of Stirling, Scotland FK9 4LA UK

Abstract-When bilaterally symmetric images are spatially filtered and thresholded, a subset of the resultant 'blobs' cluster around the axis of symmetry. Consequently, a quantitative measure of blob alignment can be used to code the degree of symmetry and to locate the axis of symmetry. Four alternative models were tested to examine which components of this scheme might be involved in human detection of symmetry. Two used a blob-alignment measure, operating on the output of either isotropic or oriented filters. The other two used similar filtering schemes, but measured symmetry by calculating the correlation of one half of the pattern with a reflection of the other. Simulations compared the effect of spatial jitter, proportion of matched to unmatched dots and width or location of embedded symmetrical regions, on models' detection of symmetry. Only the performance of the oriented filter + blob-alignmnet model was consistent with human performance in all conditions. It is concluded that the degree of feature co-alignment in the output of oriented filters is the cue used by human vision to perform these tasks. The broader computational role that feature alignment detection could play in early vision is discussed, particularly for object detection and image segmentation. In this framework, symmetry is a consequence of a more general-purpose grouping scheme.

1. INTRODUCTION

The identification of bilaterally symmetric figures is thought to be, a component of a number of basic visual tasks. It has been proposed that symmetry plays a major role in object recognition by providing a canonical axis for the representation of shape (Marr, 1982). When a symmetrical object is tilted, the degree of skew of the resulting symmetry is a cue to deriving three-dimensional structure (Kanade and Kender, 1987). It is also possible that, by identifying symmetric regions of an image, the presence of salient objects can be detected. This is' reflected by the use of symmetry, within machine-vision systems, as a cue fororienting attention (e.g. Yeshurun et al. 1992).

Barlow and Reeves (1979) demonstrated that human identification of symmetry is not an 'all-or-nothing' process but is ***graded*** with respect to a number of attributes of symmetrical patterns. They showed that discrimination of symmetrical from non-symmetrical dot textures smoothly deteriorates as the proportion of paired to unpaired

---

*To whom correspondence should be addressed.

dots in the symmetrical pattern decreases. Barlow and Reeves (1979) also showed that symmetry detection smoothly deteriorates with perturbations in dot location. They propose that human symmetry detection operates using a symmetrical distribution of feature detectors sensitive to local dot densities. Assuming that the orientation and position of the axis of symmetry are known in advance, this organisation may be practically realised by performing a simple correlation of one half of the image with a reflected version of the other. In order to account for our tolerance to spatial jitter of features, Barlow and Reeves (1979) do not correlate individual image locations, but calculate the local density of features and correlate these (coarse scale) estimates.

Jenkins (1983) noted that because a reflection is used to generate symmetrical stimuli this does not imply that a reversed mapping must be used to detect stmcture, especially given that this has no obvious interpretation in terms of the known structure of the human visual system. Instead he proposed a three-stage model of symmetry detection: detection of orientation uniformity, fusion of pairs into features, and the detection of symmetry of the resultant feature. Symmetrical dot pairs are detected directly by the output of receptive fields, of various sizes, oriented orthogonal to the axis of symmetry.

Symmetry detection seems to operate at a low spatial resolution. It is resistant to local perturbation of the position of the constituent elements of symmetrical textures (Barlow and Reeves, 1979). If symmetric textures are composed of short oriented lines, the orientation of lines does not greatly affect detection of symmetrically *positioned* features (Koeppl, 1993; Locher and Wagemans, 1993). These findings implicate mechanisms sensitive to *low spatial frequencies.* This is supported by two findings which demonstrate that the degree to which features activate filters determines their contribution to the percept of symmetry. Firstly that discrimination of symmetry from noise is not possible, at 150 ms exposure duration, when elements on either side of the axis of symmetry are of opposite contrast (Zhang, 1991). Secondly that the *proximity* of symmetrical features to one another is critical: the area around the mid-line of bilaterally symmetric texture has been found to contribute more to the perception of symmetry than other areas (Bruce and Morgan, 1975; Julesz, 1975; Barlow and Reeves, 1979; Jenkins, 1982). Violations of symmetry are more easily detected near the mid-line (Bruce and Morgan, 1975) and detection of symmetry is less affected by random displacement of dots that are nearer the outskirts than the mid-line (Barlow and Reeves, 1979; Jenkins, 1982).

These results suggest that spatial filtering-mechanisms are involved in the perception of symmetry. Recently, it has been proposed that isotropic, Laplacian-of-Gaussian (LoG) filtering might suffice for the grouping of symmetrical elements Locher and Wagemans, 1993). However, no specific scheme has been described for measuring symmetry from the output of LoGs, or from any other pre-processing system.

The starting point for our study is the observation, illustrated in Fig. 1, that when images of bilaterally symmetric objects are filtered with anisotropic filters that are oriented orthogonal to the axis of symmetry, a striking and simple pattern of response is found. In each of the filtered images in Fig. 1, there is a pattern of parallel aligned stripes in the response, each stripe centred very close to the axis of symmetry.
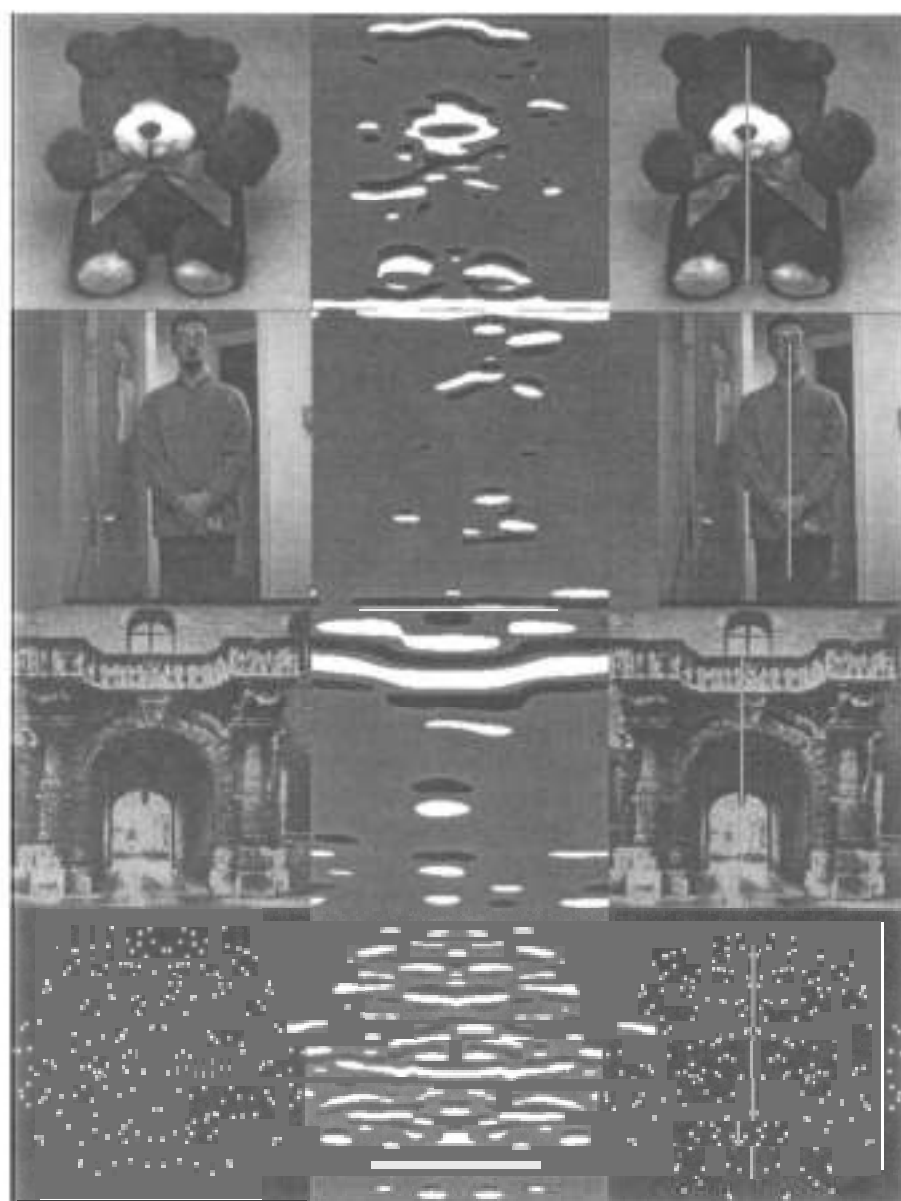
Figure 1. This figure shows (left column) natural images, (middle column) the output of a horizontal DoG at the optimal scale for symmetry detection, and (right column) the original image with the x-location which maximizes $A(x)$ marked.

Such alignment of stripes is not likely to occur by chance, and could be used as an indication of the presence of an object in the image. We have examined the hypothesis that such a pattern of response could be used by the visual system in detecting bilateral symmetry.

The proposed mechanism does not in fact detect symmetry. Bilateral symmetry is defined by the relationship:

$$I(x_a, y_a) = I(-x_a, y_a),$$

where x_a and y_a are positions orthogonal and parallel to the axis of symmetry, respectively. Any pattern that conforms to this relationship is symmetric. To measure whether a pattern is symmetric, it is strictly necessary to assess whether this relationship holds for all (x_a, y_a). A simple numerical way to perform this computation is to note that the correlation measure:

$$4 \frac{I(x_a, y_a) I(-x_a, y_a)}{\left(I(x_a, y_a) + I(-x_a, y_a)\right)^2} \tag{1}$$

reaches a maximum value of 1.0 when $I(x_a, y_a)$ equals $I(-x_a, y_a)$. This provides a means for calculating whether an image region is symmetric about a particular axis by averaging this measure for all values of $I(x_a, y_a)$. Since the symmetric pattern generally would not fill the entire image, in principle this measure would need to be assessed for all possible axes and regions. This would be achieved by applying a family of spatial weighting functions of varying size, orientation and perhaps shape, to correlations performed at every point in the image. Such a proposition would place very heavy computational demands on any neural implementation.

A much simpler mechanism, but one that does not exactly calculate symmetry, is proposed. First, we make the simplifying assumption that image values can be represented by just two values, which for convenience, we will take to be -1 and 1. We can then note that the measure:

$$I(x_a, y_a) + I(-x_a, Ya)$$

will reach an extremum value (± 2) when $I(x_a, y_a)$ equals $I(-x_a, y_a)$, and will average to the value zero otherwise. We can further note that summing along a particular row (y_a) orthogonal to the axis of symmetry:

$$\sum^{x_a} I(x_a, y_a)$$

produces a measure that will be more extreme the more similar all the values are to each other along that row. The second simplifying assumption is that where bilateral symmetry in natural images exists, it will frequently be of a form that allows this type of summation to be used. The second assumption is not as severe as it appears — given the spatial correlations that exist in natural images.

This proposal is equivalent to the starting observation, that parallel stripes occur in images that have been filtered with oriented filters. The filtering process is equivalent to some form of differentiation operation followed by an anisotropic summation. The first component will not alter bilateral symmetry, and the second component

effectively calculates the appropriate summation. Hence, any place in an image where the filtered response is high could potentially be an axis of symmetry. In practice there are many reasons why the filter response should be high. However, where several peaks in response become aligned on a common axis that is orthogonal to the preferred orientation of the filter, the likelihood is high that this identifies a region of local bilateral symmetry. Note that the computational problems associated with the correlation measure have been avoided by using the combinatorics of filtering itself.

In the study reported in this paper, implementations of the correlation measure and the alignment measure in image processing simulations of psychophysical tasks are compared with the performance of human subjects. All simulations were of the same basic form. For each condition of each task, we created a sequence of pairs of test images to be discriminated, exactly as in the real task: one member of each pair containing more symmetric structure than the other. For each pair in turn, a computational procedure was applied, resulting in a measure of degree of symmetry for each of the two stimuli. A 2AFC psychophysical response could then be generated according to the task. Four computational procedures were employed: two different types of filter (isotropic and oriented) and the two different types of symmetry measure (correlation and alignment).

2. GENERAL METHODS

2. 1. *Filtering*

There is accumulating evidence that the early visual system consists of mechanisms selective for certain spatial frequencies in the input (e.g. Campbell and Robson, 1968; Sachs *et al.*, 1971). Models that have been proposed to describe the exact point-spread function of cells in visual cortex include Gabors (e.g. Daugman, 1985) and Difference-of-Gaussians (DoGS) (e.g. Wilson and Gelb, 1984). Similarly, it has been proposed that retinal ganglion or LGN cells can be described by the Laplacian-of-Gaussian (e.g. Marr, 1976). The principal functional difference between these models is whether filters are oriented or isotropic. In the present study both isotropic and oriented filters are examined.

The isotropic filter used was the Laplacian-of-Gaussian with a point-spread function defined as:

$$f(x, y, s) = \frac{1}{s^2}\left(1 - \frac{x^2 + y^2}{s^2}\right)e^{-(x^2+y^2)/2s^2},$$

where s is the space constant of the filter.

For the oriented filter the point-spread function used was an elongated, horizontal DOG filter:

$$f(x, y, s) = \left(e^{-y^2/2s^2} - \frac{1}{2.23}e^{-y^2/2(2.23s)^2}\right)e^{-x^2/2(3s)^2},$$

where s is the space constant of the filter. The ratio of the amplitudes of the positive and negative parts of the DoG and the aspect ratio of the filter are based on those derived by Wilson and co-workers (Phillips and Wilson, 1983; Wilson and Gelb, 1984). Because only vertical bilateral symmetry is under consideration, only horizontally oriented filters are considered.

Seven sizes of LoGs were used: s = 2.00-16.00 pixels, in multiplicative steps of $\sqrt{2}$. Seven sizes of horizontal DoG filter were used: $s = \sqrt{2}-8\sqrt{2}$ pixels, in multiplicative steps of $\sqrt{2}$. Both sets of filters had peak spatial-frequency sensitivities of between 28.8 and 3.5 cycles per image, respectively.

2.2. *Primitive extraction*

The result of the filtering is divided into positive and negative signals, by thresholding the image. Grey levels less than one standard deviation greater or less than the mean grey level are set to zero. This non-linearity is introduced as a way of delineating individual image features.

Filter output images are then converted into a symbolic or 'primal sketch' type representation (Marr, 1976). Blobs from the filtering/thresholding stage are described using measurements of their centroids, lengths, etc. Such schemes have previously been proposed as practical methods for deriving texture statistics (e.g. Voorhees and Poggio, 1987). Watt's (1991) image description scheme was used to describe each zero-bounded region blob in the form:

$$(\text{cx}, \text{cy}, \mu, \lambda, \theta),$$

where (cx, cy) is the centroid, μ is the mass, λ is the length, and θ is the orientation of the blob. (Details of how these parameters are derived are described in Watt (1991), pp. 114-120.)

2.3. *Measuring alignment*

Consider the alignment, *A(x)*, of all blobs which intersect a particular image column, x. A simple measure of blob alignment is to calculate the distance from x to the centroid of the blob:

$$A(x) = \frac{1}{M} \sum_{i=1}^{N_x} \exp\left[-\left(\frac{x^2 - cx_i^2}{2\lambda_i^2} \right) \right] \mu_i, \qquad (2)$$

where N_x is the number of blobs intersecting the xth image column, and M is their total mass. Deviation of the blob centroid from x is weighted in inverse proportion to the length of the blob, λ_i, and in direct proportion to the mass of the blob, μ_i. Normalization by M means the alignment measure falls in the range 0.0-1.0. The maximum value of A for the image, $\max_x(A)$, is used as a measure of the symmetry of the pattern. The location at which the maximum occurs x_{max} is used to indicate the axis of symmetry. Figure 1 demonstrates the application of this technique to axis location in natural and artificial symmetrical images.

2.4. Correlation

Equation (1) was used to calculate an image correlation for each image.

2.5. Estimating discriminability

Given two sets of measurements from each model, one of the reference set and one of a particular level of the cue, these sets will have means of μ_{ref} and μ_{cue}, respectively, and standard deviations of σ_{ref} and σ_{cue}. Typically, the cue and reference sets had unequal variance, suggesting an appropriate criterion as:

$$c = \mu_{ref} + \frac{\sigma_{ref}}{\sigma_{ref} + \sigma_{cue}} (\text{h e } - \mu_{ref}). \tag{3}$$

Values from the reference and cue files were randomly selected and compared to the criterion to establish probability of detecting a cued stimulus (hit), versus probability of falsely identifying a reference stimulus as the cue (false alarm). From these two values d' was calculated as:

$$d' = P^{-1}(\text{hit}) - P^{-1}(\text{false alarm}),$$

where $P^{-1}(y)$ is the inverse function of the Gaussian probability function:

$$P(y) = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{Y} e^{(-x^2/2)} dx.$$

This function can be calculated to an arbitrary level of precision using an approximation to the incomplete Gamma function (Press *et al.*, 1992).

3. SIMULATION PROCEDURE

For each image, symmetry measures were derived using four models. The first two models used the alignment measure in conjunction with oriented or isotropic filters (the 'I + A' and '0 + A' models, respectively). Alignment estimates were made, as described above, and these values placed in files.

The third and fourth models were correlators, operating on the output of similar oriented and isotropic filters. The convolved images were divided at the axis of symmetry, correlated with a reflected version of the other half, and the degree of correlation recorded. These are referred to as the 'I + C' and the '0 + C' models, according to the type of pre-processing used.

All models had prior knowledge of the location of the axis of symmetry. This meant that the degree of alignment or correlation was calculated only around the central image column. Uncertainty on axis location/orientation is known to reduce performance of human subjects on this task (Barlow and Reeves, 1979) and similarly affects the output of the models.

4. SIMULATIONS 1 AND 2 EFFECT OF SIGNAL-TO-NOISE RATIO AND POSITIONAL JITTER

Barlow and Reeves (1979) measured the discriminability of symmetric from non-symmetric random-dot textures as a function of the ratio of paired to unpaired dots in the symmetric texture. The ratio of paired to unpaired dots (the signal-to-noise ratio, SNR) was varied and discrimination from pure noise and pure symmetry was measured. As the number of paired dots in the stimulus decreases (a) discrimination from noise deteriorates and (b) discrimination from pure symmetry improves.

Barlow and Reeves (1979) also showed that discrimination of symmetry from noise deteriorates as jitter on the position of symmetrical elements increases. The second simulation in this section is to investigate if operation of the models at coarse spatial scales can explain subjects' performance in this condition.

4. 1 *Stimuli*

The stimuli were textures composed of 100 dots spatially randomly distributed in a centrally-positioned circular region, with radius 128 pixels. Each dot was composed of a square of four individual pixels and appeared black on a white background.

In generating noise stimuli, dots were allowed to fall randomly in the circular field. The symmetrical patterns were generated by dropping a fixed proportion of dots randomly in the circular field and then reflecting each of their positions around a vertical, centrally positioned axis of symmetry and placing a corresponding dot there. In the second simulation, dot locations were generated as above and then each was subjected to a random shift.

4.2. *Procedure*

For Simulation 1, 128 stimuli were generated at each of eleven ratios of paired to unpaired elements (0.0-1.0, in steps of 0.1). For Simulation 2, 128 stimuli were generated with between ± 4.06 and ± 65.02 pixels of positional jitter in multiplicative steps of $\sqrt{2}$ pixels (corresponding to shifts of ± 2 to ± 32 arcmin, at the viewing distance quoted in the original paper). For each stimulus level, and for each model, d' was measured in a procedure exactly analogous to that employed in the original experiments.

4.3. *Simulation results and discussion*

Results from Simulation 1 are presented in Fig. 2. The left half of the figure shows results for discrimination of an increasing proportion of symmetry from noise. In accord with human data, all four models predict a gradual improvement in performance with increasing proportion of symmetry. As the spatial scale is increased, the predictions of the 0 + C model match human data more closely. Note that both of the alignment models are less sensitive to the scale of filter selected.

The right half of the figure shows the predicted discriminability of pure symmetry from increasing proportions of symmetry. The correlation models overestimate performance, typically because their estimates of *pure* symmetry show little or no variation
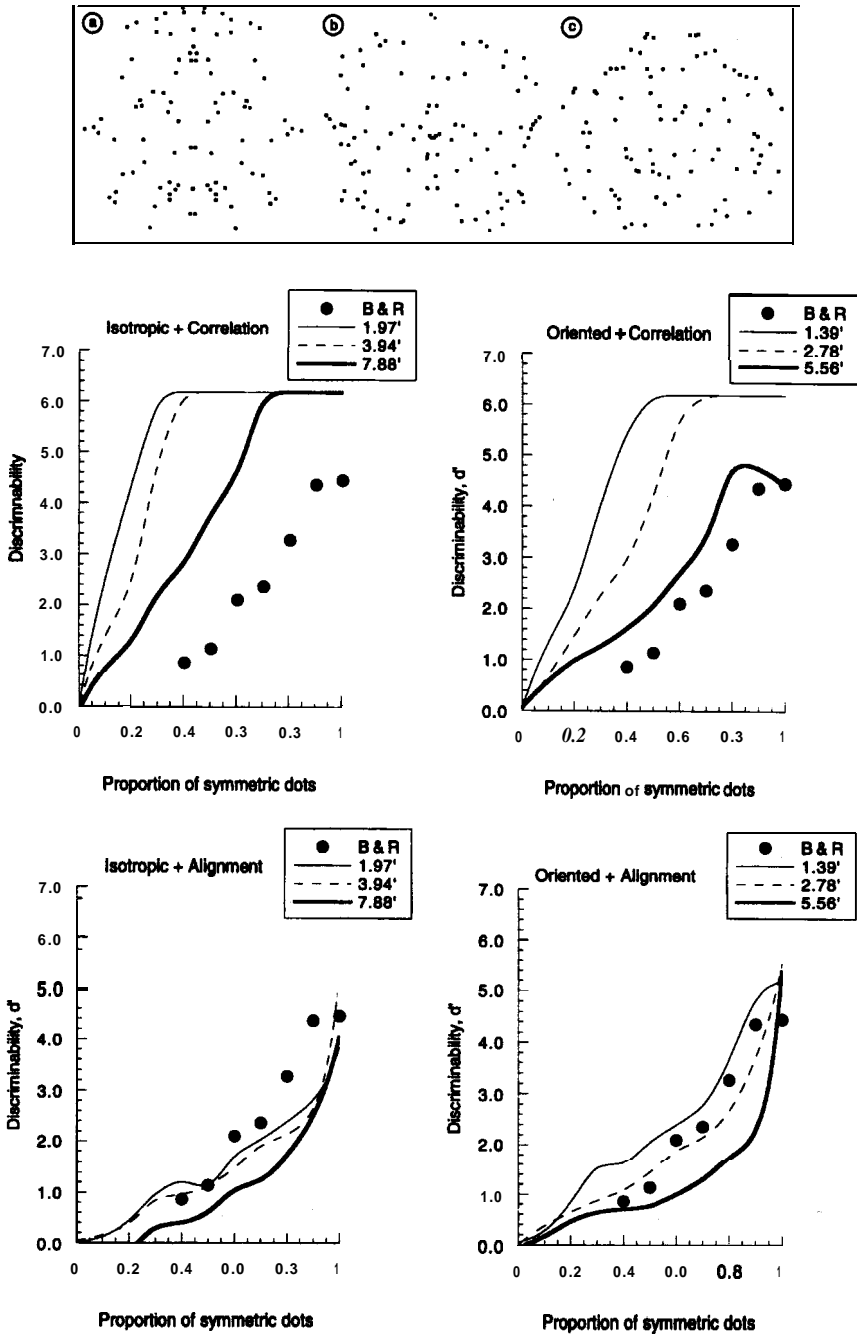
Figure 2. (Top part) Examples of the stimuli used in Simulation 1. Patterns have the following proportions of paired dots: (a) 1.0, (b) 0.75, and (c) 0.5. (Bottom part) Solid symbols show data from Barlow and Reeves (1979), for discrimination from (this page) noise and (next page) pure symmetry patterns, as a function of the proportion of paired to unpaired dots in the stimulus. Lines show predictions from models using filters with space constants given in the legend (in arcmin).
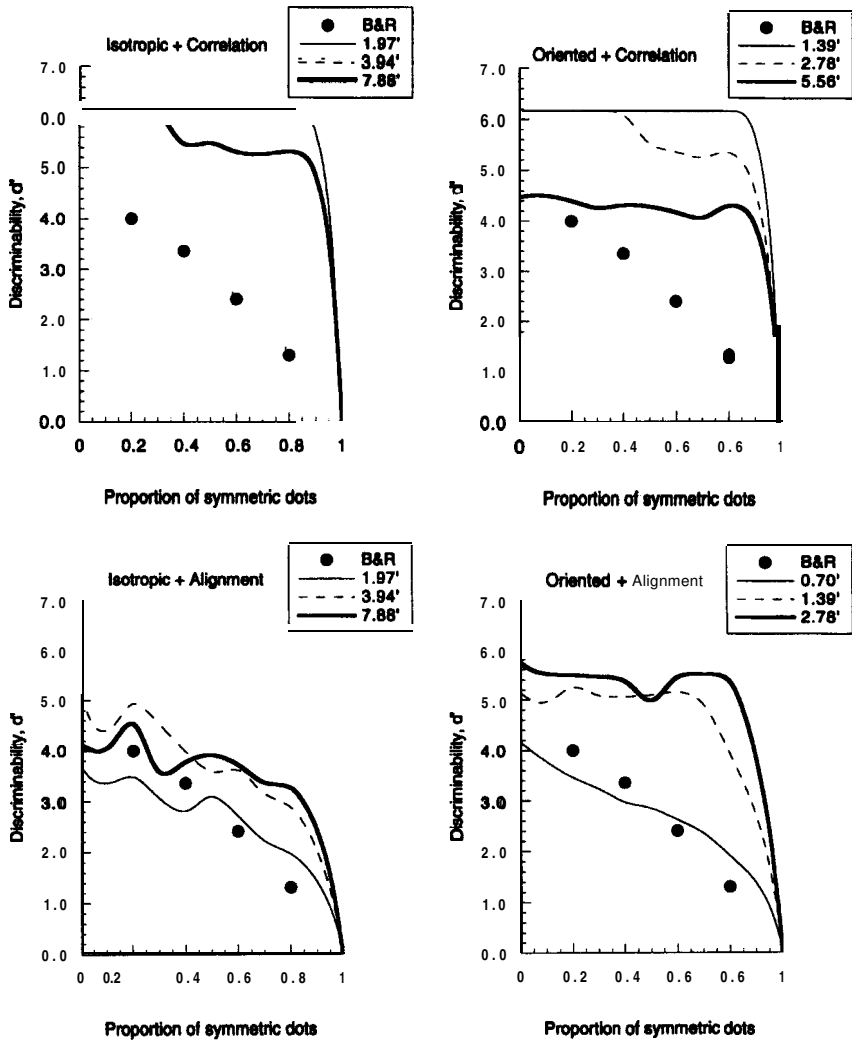
Figure *2.* (Continued).

(around 1.0), so d' measures tend to be very high. Predictions from the alignment models, however, do show the same gradual deterioration as human subjects. These models introduce variability in the representation of pure symmetry that is consistent with human performance on this task.

Simulation results for structure discrimination in the presence of positional jitter are plotted alongside human data in Fig. 3. All models produce a gradual drop-off in discrimination of symmetry from noise, as positional jitter increases. However, both correlation models overestimate performance at very low levels of jitter. The I + A model fails to achieve human levels of performance in all but the extremely perturbed patterns. A failure to fit because of overly high performance can be accounted for by the assumption of additional noise in a system, but a failure to achieve human
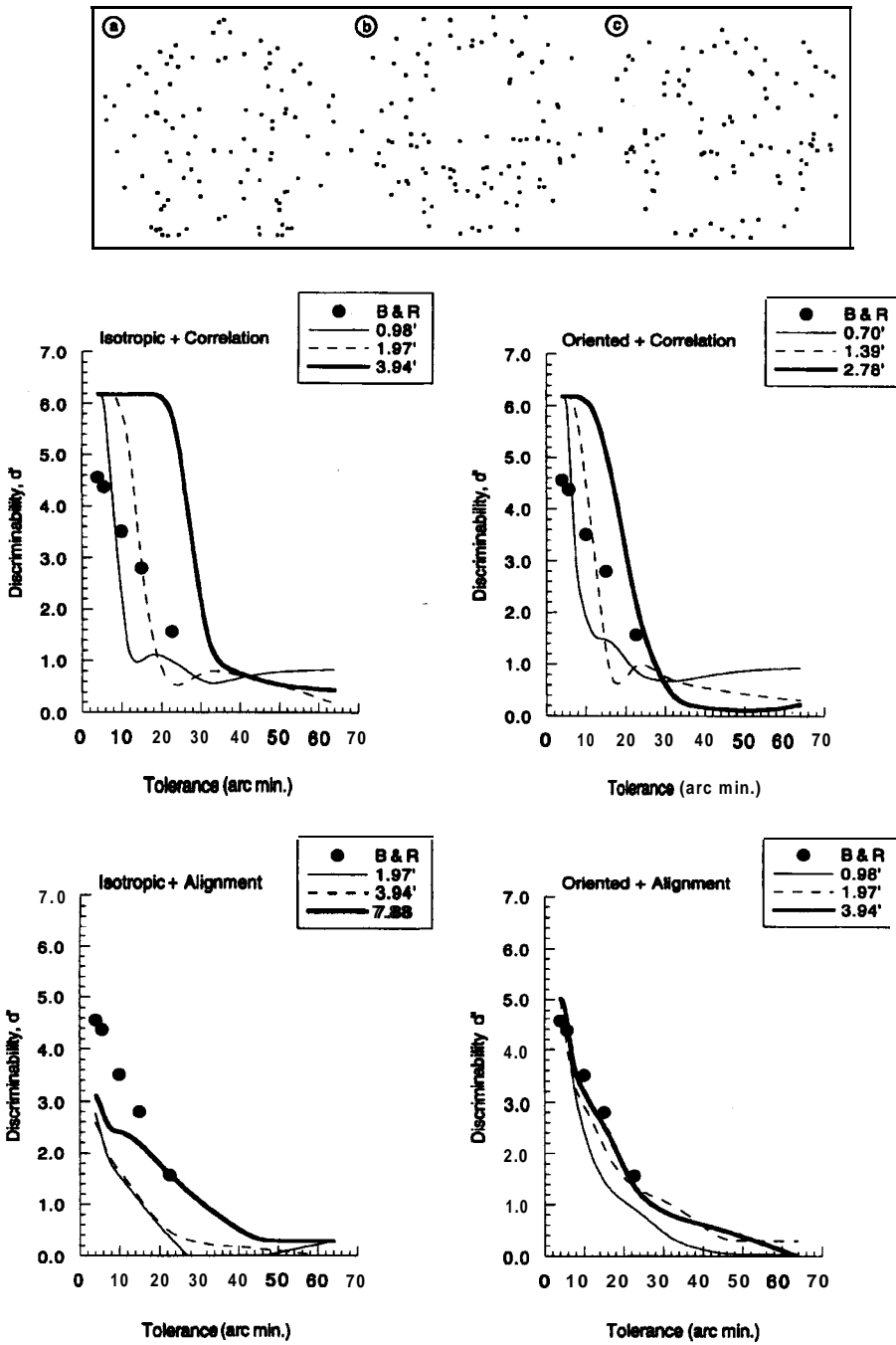
Figure 3. (Top part) Examples of the stimuli used in Simulation 2. Patterns have positional jitter of (a) ±4.0, (b) ±8.0, and (c) ±16 pixels. (Bottom part) Solid symbols are psychophysical data from Barlow and Reeves (1979), for discriminating symmetrical patterns from noise in the presence of spatial jitter. Lines represent the predictions from the models at a number of spatial scales.

levels of performance must be taken as very strong evidence against the model. The agreement of the 0 + A with human data models is extremely good, especially given that no fitting of the model to the data has been used. Note that the 0 + A model predicts similar performance across two octaves of filter sizes.

All of the proposed models produce a graded response to symmetry in the presence of unpaired dots, but the I + C model, consistently overestimates performance. In the presence of spatial perturbation of elements, the I + A model fails to perform as well as humans and for that reason can be rejected. The correlators' behaviour is broadly in accord with human data, but the fits of the 0 + A model are consistently better.

5. SIMULATIONS 3-5: EFFECT OF LOCATION OF SYMMETRY WITHIN A TEXTURE

A pattern containing a set of symmetrical dots embedded in a background composed of noise dots is most effectively discriminated from noise patterns when the symmetrical region is located around the axis of symmetry (Barlow and Reeves, 1979). However, the next most discriminable region is at the outer boundaries of the pattern. The region producing lowest discriminability is located between the axis and boundary. If any model using filtering predicts that symmetry is simply related to the proximity of the closest matched pairs then there should be no such advantage for structure around the boundary. In this section we simulate three psychophysical tasks examining the role of the *locution* of the symmetrical features within the pattern.

Jenkins (1 983) measured the effect of the spatial location of a symmetrical region embedded in noise. The tasks measured discriminability of patterns (from noise or pure symmetry) as a function of the width of the embedded symmetrical region. The tasks were discrimination of: (a) noise patterns from noise containing symmetry around the axis, (b) perfect symmetry from symmetrical patterns with noise encroaching from the boundary, and (c) noise patterns from symmetrical patterns containing noise around the axis of symmetry.

5.1. Stimuli

Stimuli were generated in approximate accord with those described in Jenkins (1983).[1] Textures contained 650 dots, where each dot was an individual pixel, appearing white on a black background.

The stimuli for Simulations 3 and 4 were noise textures with a strip of symmetrical dots around the axis of symmetry. Dots were randomly positioned throughout the image, except for a region around the axis of symmetry. A proportion of the total number of dots was then placed in the central strip, and reflected around the axis. The number of dots in the central region was proportional to the area of the strip, so that there was no difference in density across the pattern. The reference stimuli for Simulation 3 were noise textures (i.e. the task was to detect the pattern which contained a strip of symmetry), and the references for Simulation 4 were purely symmetrical patterns (i.e. the task was to detect which pattern had noise encroaching from the boundaries).

The stimuli for Simulation 5 were generated in a similar way to above, except that textures were symmetrical but contained a strip of randomly positioned elements around the axis of symmetry. The reference stimuli used were pure noise patterns (i.e. the task was to detect which patterns had symmetry at the boundaries).

5.2. *Procedure*

128 stimuli were generated at each width of symmetrical/noise strip. The width of strips used were: for Simulation 3, 0-24 pixels, in steps of 4 pixels (i.e. from O-O.468 deg. assuming a 5 deg wide display), for Simulation 4, 0-72 pixels, in steps of 8 pixels (i.e. from O-l.41 deg), and for Simulation 5, 0–24 pixels, in steps of 4 pixels. Measurements of symmetry were made using the same models described in the previous simulation, and percent correct results were generated as in Jenkins (1983).

5.3. *Simulation results and discussion*

Figure 4 shows that humans show a steady increase in the discrimination of an expanding symmetrical strip from noise, as the width of symmetrical strip increases, with performance levelling off at around 0.22 deg. All of the models approximately show this behaviour but their predictions do not deteriorate as sharply as human data: humans are very poor when strip width falls to around 0.1 deg. This discrepancy is probably due to two characteristics of the alignment models' operation. Firstly it is noise-free with respect to the number and position of blobs. One might expect that smaller samples of tiny blobs would give worse estimates of alignment, either because blobs would fail to be registered, or because any noise on the alignment measurement would become large in relation to the width of the blobs. Secondly, there is no uncertainty of the location of the axis — all models know exactly which image column to use. Human subjects, even given that the axis did not shift between stimuli, probably have some error on locating the axis from trial-to-trial.

Figure 5 shows that as the width of a symmetrical strip increases (to a maximum value of only 20% of the total pattern width) subjects find it increasingly hard to discriminate patterns from pure symmetry. At the point where subjects can no longer discriminate, however, 80% of the pattern is symmetrical. Consequently, the correlators predict perfect discrimination for the widths tested. The alignment models, on the other hand, predict the steady decrease in discrimination performance. The fit of the 0 + A model is the closest.

The final simulation was of human discrimination of symmetrical textures containing strips of noise around the axis, from noise textures (Fig. 6). Our ability to perform this discrimination deteriorates with increasing width of the strip, and at a noise width of around 1.0 arcmin subjects are approaching chance performance. Again the correlators predict no change in performance because patterns contain so much symmetry. The alignment models' sensitivity to the *location* of the symmetry is again shown in the collapse of performance as noise width increases. Interestingly, however, the I + A model collapses too quickly; isotropic filtering mechanisms are insufficiently sensitive to the structure when the noise strip becomes wide. The
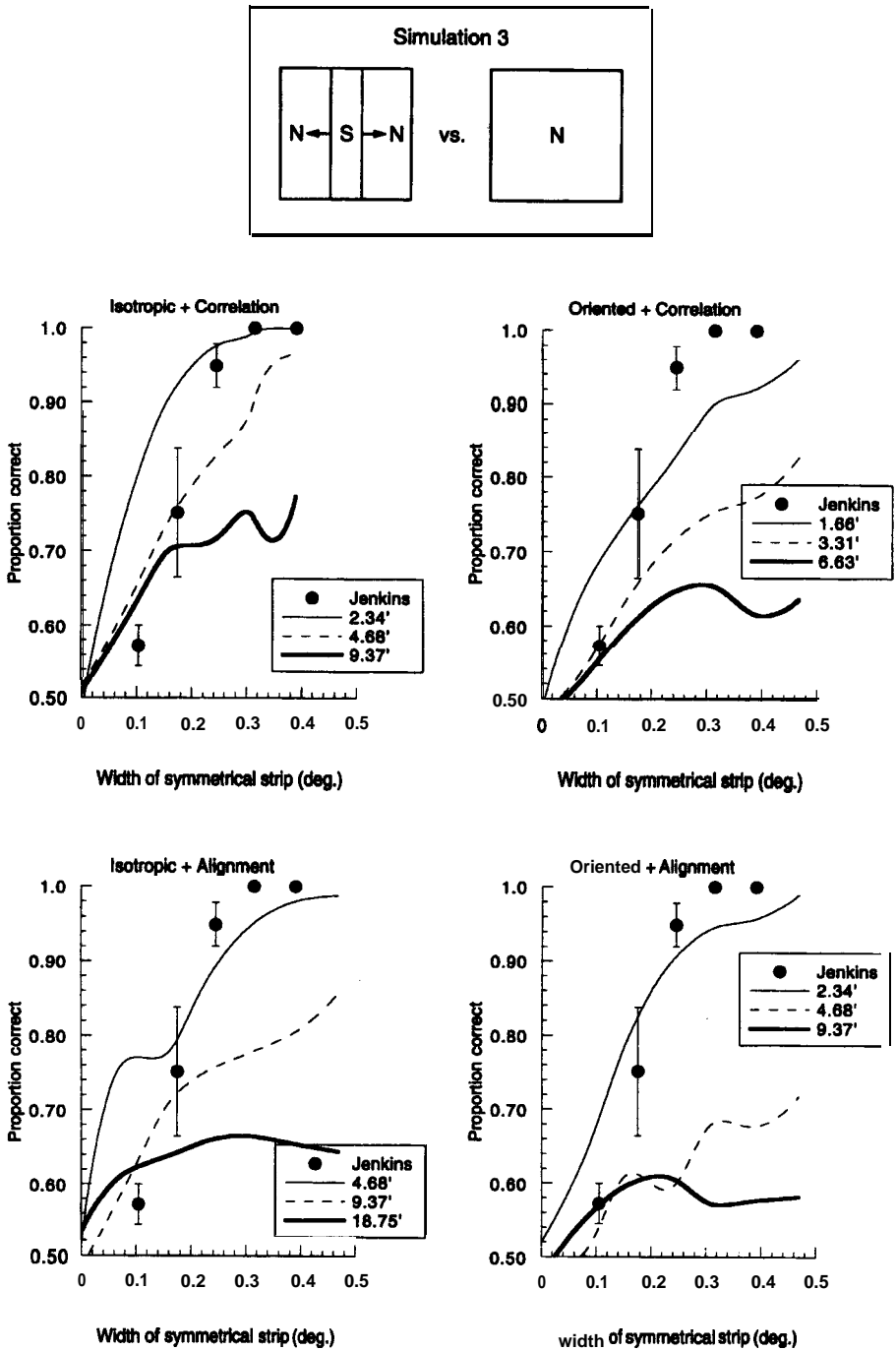
Figure 4. (Top part) Schematic representation of the task in Simulation 3. 'S' indicates symmetry, 'N' represents noise. (Bottom part) Solid symbols show human discrimination of such textures from pure noise. All models produce improved behaviour as a function of the width of the band.
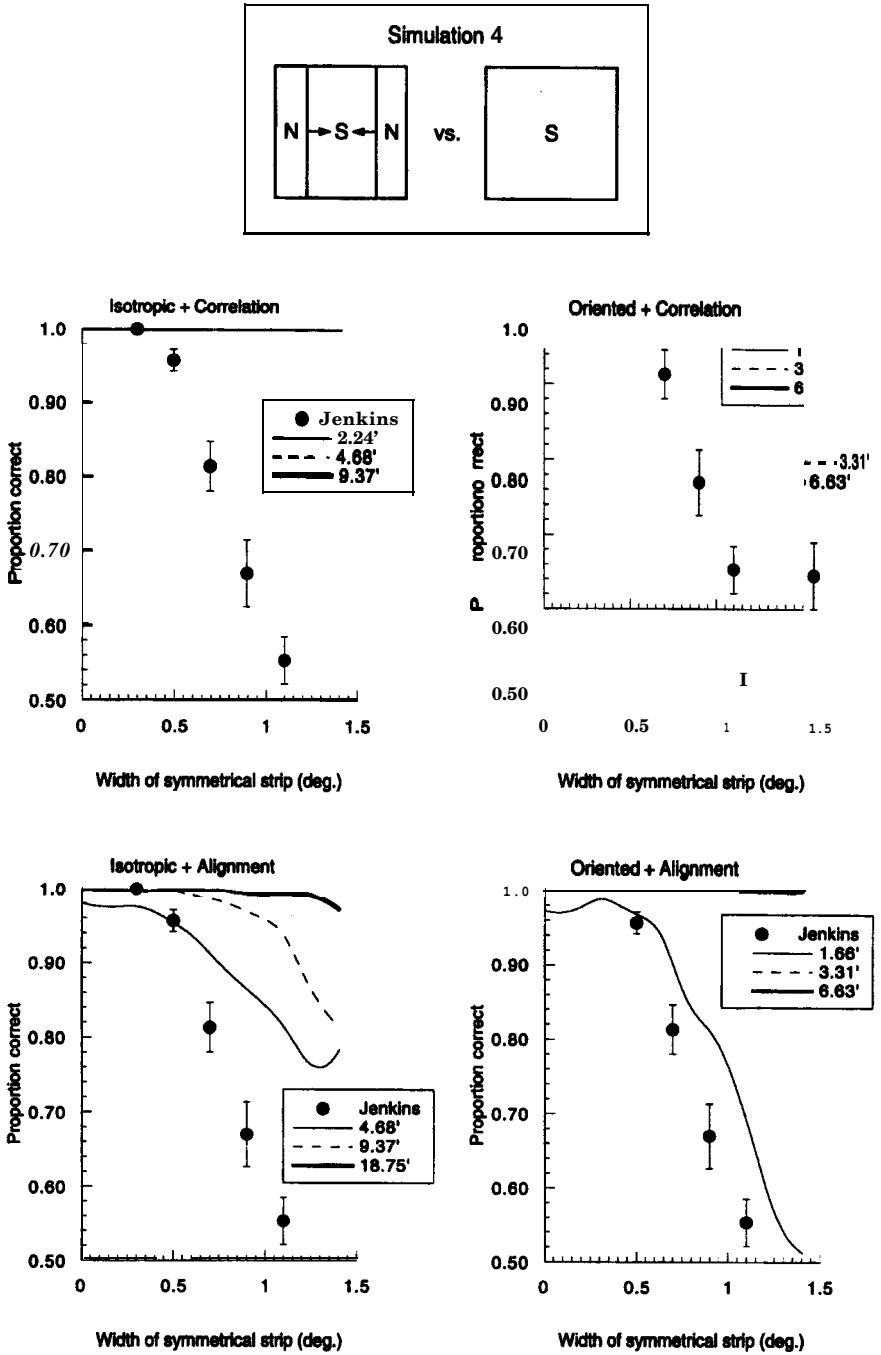
Figure 5. (Top part) Schematic **representation** of the task in **Simulation 4.** (Bottom part) **Human per-**formance for **discrimination** of an **embedded** band of symmetry from pure symmetry is **represented** by solid symbols. Only the performance of the alignment models show a dependence on the widtb of the symmetrical strip.
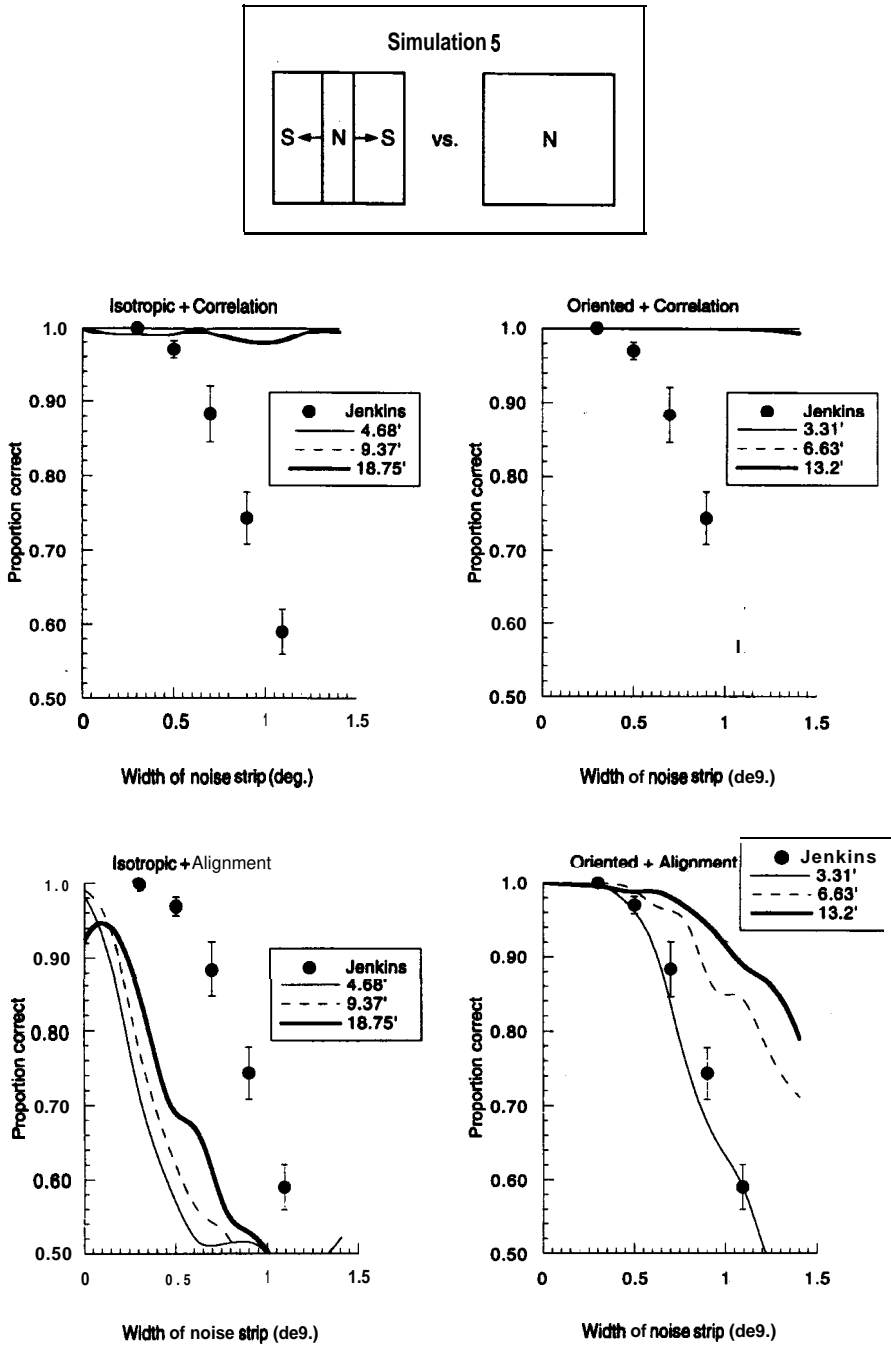
Figure 6. (Top part) Schematic representation of the task in Simulation 5. (Bottom part) Human discrimination of such textures from pure noise. Although all models are capable of detecting symmetry in these patterns, only the alignment models' performance breaks down, with increasing width of the noise strip, in accord with human data.

0 + A model captures the best of both models. It can overperform, but certain scales (s.d. = 3.31 arcmin) produce patterns of discrimination which match human data. It is interesting, and somewhat counterintuitive, that an alignment model can detect correlation when the area around the axis of symmetry is pure noise. Note however that the filter size producing the best predictions are larger than those used in the previous two simulations. This also suggests an explanation for why structure at the periphery is more efficiently detected than structure embedded between axis and boundary (Barlow and Reeves, 1979). Small filters are responsible for detecting structure at the axis, large filters are responsible for structure at the boundaries. (We return to this point in the Conclusion section.)

To summarise: results from Simulation 4 indicated that correlators overestimate discrimination of symmetrical strips embedded in noise from pure symmetry. Alignment models account for human data well. Simulation 5 showed that only the 0 + A model could account for the deterioration shown by human subjects in discriminating symmetrical textures containing an expanding strip of noise around the axis from noise. The correlators again systematically overestimate, and the I + A model systematically under-estimates, human performance. Only the 0 + A model provides consistently good fits to the data across all simulations.

6. CONCLUSIONS

The results of the simulations presented in this paper lead to a number of tentative conclusions. First, we have shown the adequacy of the alignment measure for detecting symmetry. Second, we have shown that a model based on alignment but using an isotropic filter is inadequate, especially since it fails to reach the human levels of performance. Third, we have shown that the correlation measure of symmetry does not match human data closely with either form of filter, in several cases being much better in its performance than are the subjects. Fourth, we have shown that an alignment measure, after filtering with oriented filters, produces data that is in close agreement with subjects for all tasks considered.

For each simulation, results from a range of different spatial scales are reported. In every case the output of a single filter is used to reach a psychophysical decision and hence a psychometric function. It might be thought that combining information from the outputs of different filters would be a way of improving performance, for example through probability summation. In practice, this is not found to be the case. The reasoning is as follows. In reaching a psychophysical decision, the main causes of an incorrect response are 'spurious' near-symmetry responses in the noise targets. These are most awkward in filters that are not well matched, spatially, to the spatial structure of the target. Thus combining responses from different filters increases the probability of detecting the symmetry, but also increases the probability of detecting spurious symmetry. Probability summation therefore would not improve performance.

To conclude, we shall discuss why the alignment model with oriented filters succeeds when the others do not and then we shall go on to place the model in the context of a general-purpose visual process. Our claim is that symmetry detection, when
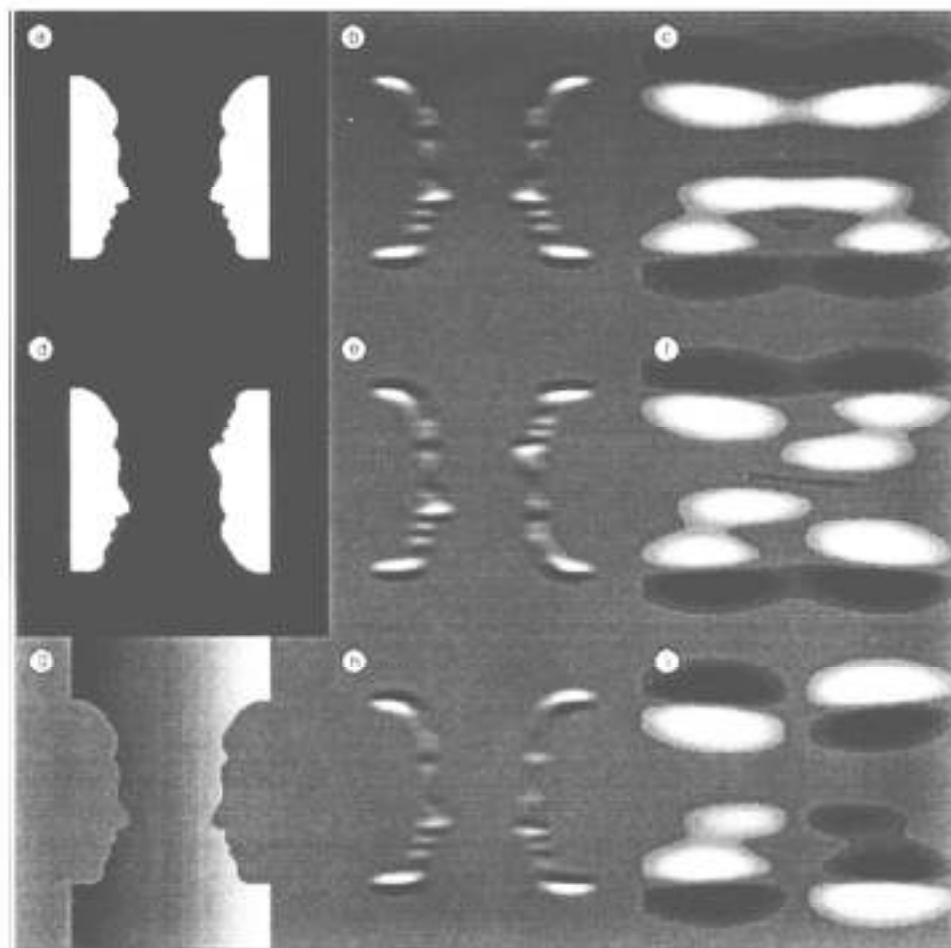
Figure 7. (First column) Variants on the Rubin vase: (a) standard vase, (d) figure with one inverted face, (g) figure with vase shaded by an intensity gradient. (Second column) Fine scale. (Third column) Coarse scale horizontally filtered versions. Note the clustering of blobs around the facial profiles in the second column, and the alignment of blobs around the centre of the symmetrical vase in (c).

conceived of as being performed by the alignment model, is a special case of a more general, object-detection mechanism.

First, we consider why the alignment model, with oriented filters, does match human performance. This model is based on the proposition that parallel and aligned correlations in images are worth detecting, a point we return to below. In practice, most randomly produced bilaterally symmetric patterns have structure of this type. They need not; for example, a checkerboard pattern will have bilateral symmetry that lacks continuous correlations. The critical data here, are the effects of which parts of the target are symmetric. Given the need for continuous correlations, it is not surprising that the alignment model detects symmetry about a mid-line efficiently, but does not detect symmetry when it does not extend over a sufficiently large area around the

mid-line. A correlation model, on the other hand, has no specific requirement for the correlation to extend through the mid-line region. The alignment model only works with oriented filters. Isotropic filters do not produce extended regions of response.

For most of the simulations that we have run, the choice of spatial scale for the oriented filter is not critical for the alignment model. In general, human data match the performance of a fairly fine-scale filter. For some simulations (1, 2), the choice of filter spatial scale barely affects model performance. In Simulations 3 and 4, subjects match the performance of fine-scale filters. In Simulation 5, subjects match the performance of a slightly coarser scale filter. There is thus no one spatial scale at which human data can be uniformly modelled, although a parsimonious account would be that fine-scale filters ($s \approx 1.5$ arcmin) are used in all cases except Simulation 5. The cue in Simulation 5 is structure that is further from the mid-line of the pattern, and it is perhaps not surprising that a coarser scale serves best here. A similar point is discussed below and illustrated in Fig. 7.

Now, we consider whether the alignment model is plausible in the context of a visual system. The main components of the model are oriented filters and an alignment detection mechanism. Oriented filters are known to exist in visual cortex, but to our knowledge an alignment mechanism has not been sought, although the 'non-Cartesian' receptive fields reported by Gallant *et al.* (1993) may be suitable candidates. In this respect, it is important to note that a simple linear receptive-field type of weighting function will not suffice to detect alignments. Any mechanism that serves the function we propose must be broadly insensitive to the separation between the aligned stripes and to the sequence of their polarities. Since varying the separation and polarity of the aligned stripes would lead to variations in the spatial-frequency amplitude and phase contents of the pattern, such a requirement would need a receptive field that was broadly tuned and had complex-cell types of properties.

Our final claim is that the proposed alignment mechanism is a general-purpose strategy for detecting objects, or salient structures, in images and that symmetry is a special case of this. Figure 7 shows an example of the logic of our argument. The first column of the figure shows three variants of the Rubin vase figure. The first image shows a normal version; the second contains an inverted face profile; the third, a version where the 'vase profile' is shaded with a continuous intensity gradient.
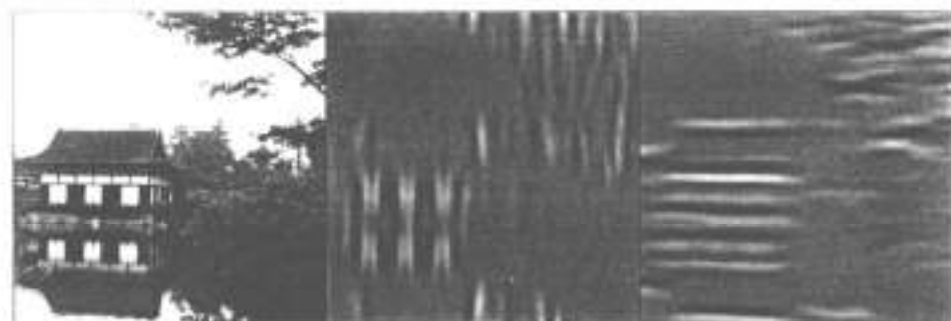

Figure 8. A scene containing natural and man-made objects, and versions filtered with horizontal and vertical filters.

The second and third columns show fine and coarse-scale filtered versions of each image. There are two types of alignment pattern revealed in the filtered images: one is related to the 'vase profile' (Fig. 7c) and one is related to the 'face profiles' (Fig. 7b, e, h). The vase case can be regarded as representing symmetry detection, but the second cannot. They both, however, can be regarded as alignment detection. In this sense, symmetry can be seen to be a special case of a more general-purpose structure detection.

Finally, Fig. 8 shows the results of filtering a real image with horizontal and vertical oriented filters. Note that the responses have aligned stripes along salient objects, man-made and natural. This point is elaborated in Watt (1994).

Acknowledgements

NOTES

1. Jenkins used *dynamic* dot displays. Each dot was sequentially generated (every $12\frac{2}{\mu}s$) and very briefly displayed ($1.5\mu s$). Thus 16446 points per sec were generated and plotted although, at any one moment, there was only one dot pair on the screen. In the simulation described, static patterns were used and matched for *perceived* density: reported as 26 points per deg^2.

REFERENCES

Barlow, H. and Reeves, B. (1979). The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vision Res.* 19, 783-793.

Bruce, V. and Morgan, M. (1975). Violations of symmetry and repetition in visual patterns. *Perception* **4**, 239-249.

Campbell, F. and Robson, J. (1968). Application of Fourier analysis to the visibility of gratings. *J. Physiol.* 197, 551-566.

Daugman, J. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimised by two-dimensional cortical filters. *J. Opt. Soc. Am.* 2, 1160– 1169.

Gallant, J., Braun, J. and Van Essen, D. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* **259**, 100– 103.

Jenkins, B. (1982). Redundancy in the perception of bilateral symmetry in dot textures. *Percept. Psychophys.* 32, 171-177.

Jenkins, B. (1983). Component processes in the perception of bilaterally symmetric dot patterns. *Percept. Psychophys. 34,* 433–440.

Julesz, B. (1975). Experiments in the visual perception of texture. *Scient. Am.* 232, 34-43.

Kanade, T. and Kender, J. (1987). Mapping image properties into shape constraints: Skewed symmetry, affine-transformable patterns, and the shape-from-texture paradigm. In: *Readings in Computer Vision: Issues, Problems, Principles, and Paradigms.* M. Fischler and O. Firschein (Eds). Morgan Kaufmann, California, pp. 257-267.

Koeppl, U. (1993). Local orientation versus local position as determinants of perceived symmetry. *Perception 22* (Suppl.), 111.

Locher, P. and Wagemans, J. (1993). Effects of element type and spatial grouping on symmetry detection. *Perception 22, 565-587.*

Marr, D. (1976). Early processing of visual information. *Proc. Roy. Soc. Lond.* B 275, 483-534.

Marr, D. (1982). *Vision.* Freeman, San Francisco, CA.

Phillips, G. and Wilson, H. (1983). Orientation bandwidths of spatial mechanisms measured by masking, J. *Opt. Soc. Am.* **Al, 226-232.**

Press, W., Teukolsky, S., Vetterling, W. and Flannery, B. (1992). *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, Cambridge.

Sachs, M., Nachmias, J. and Robson, J. (1971). Spatial frequency channels in human vision *J. Opt. Soc. Am.* 61, 1176– 1186.

Voorhees, H. and Poggio, T. (1987). Detecting textons and texture boundaries in natural images. In: *Proceedings of the First International Conference on Computer Vision.* pp. 250–258.

Watt, R. (1991). *Understanding Vision.* Academic Press, London.

Watt, R. (1994). A computational examination of image segmentation and the initial stages of human vision. *Perception* (In press).

Wilson, H. and Gelb, D. (1984). Modified line-element theory for spatial-frequency and width discrimination. *J. Opt. Soc. Am.* **Al,** 124– 13 1.

Yeshurun, Y., Reisfeld, D. and Wolfson, H. (1992). Symmetry: A context free cue for foveated vision. In: *Neural Networks for Perception Vol. I.* H. Wechsler (Ed.). Academic Press, London, pp. 477–491.

Zhang, L. (1991). *Symmetry perception in human vision.* PhD Thesis, University of Trieste.