NASA Technical Memorandum 84353

Detection and Recognition of Simple Spatial Forms

Andrew B. Watson

April 1983

National Aeronautics and
Space Administration

NASA Technical Memorandum 84353

Detection and Recognition of Simple Spatial Forms

Andrew B. Watson, Ames Research Center, Moffett Field, California

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field. California 94035

DETECTION AND RECOGNITION OF SIMPLE SPATIAL FORMS.

Andrew B. Watson
NASA Ames Research Center
Moffett Field, CA 94035, USA

1. Models in Vision

In spite of the large number of intelligent, energetic people engaged in the study of vision, it often seems that we advance our understanding of this process at an excruciatingly slow pace. Why is this so? I believe that it is the fault of our models. We have been seduced by the simplicity of explanation available in the physical sciences, and try to describe the vast complexity of vision with models that would not do justice to a sphere rolling down an inclined plane.

Today I will describe a model which in a small way attempts to remedy this situation. The principles that underly it are simple, but it attempts to represent more adequately the inherent complexity of vision. Before describing this model in detail, I will emphasize its unusual aspects, and some of the benefits we may gain from this departure.

First, the model is explicit. No "channels", "mechanisms" or other ill-defined entities appear. Put another way, the model is computable. It is well enough defined to permit numerical predictions to be made for a given experimental situation. Second, the model adequately represents all the dimensions that are pertinent to the experimental domain. In the case I consider, these are the two spatial dimensions of a monochromatic, stationary image. Although this is an inherently two-dimensional (2D) situation, visual models in this context have almost invariably been one-dimensional (1D). Third, the model is general within its specified domain. The same model can be used to predict performance in a wide variety of different of visual tasks on a wide variety of images. This contrasts with most visual models, which only attempt to explain the data from a very restricted domain. Fourth, the model draws many of its parameters and assumptions directly from physiological results. While this has long been an ambition of psychophysical theories, it has rarely been realized. Finally, the model attributes intelligence to the observer. It is perhaps understandable that this has not been done in the past, since psychophysicists are most often their own observers. The model assumes that the observer will make optimal use of the available information. This assumption is of course a commonplace in signal detection theory [10], but rarely appears in explicitly visual models.

The price of these amendments is a model that is somewhat more complex, requiring more assumptions and parameters than average. The benefits, however, are worth it. It provides a mechanism with which to integrate information within the field. Since the model can be applied in diverse contexts, it provides a common repository for results from many different sources. The model also insures the consistency of interpretation from one experiment to the next. Too often a model is constructed in one context that is quite obviously incompatible with data from another context. The model also provides a natural path along which to specify in ever greater detail the relation between our visual experience and the physiological mechanisms of the visual brain. Finally, to the extent that the model is successful, it will permit us to turn our attention from the early, image-driven stages of vision to the more complex, cognitive processing that must subsequently occur.

It should be clear that the argument I have advanced is on behalf of any model that meets the

criteria of explicitness, adequacy, and generality noted above. The model I will describe is just a candidate, and a green one at that.

2. Domain of the model

What is the domain of my model? It is intended to account for psychophysical responses of an accommodated, fixating human observer viewing binocularly a 2D stationary achromatic image at a fixed adapting level of 340 cd m^{-2}. Specifically, I have excluded both color, stereo, and dynamic imagery since it seems likely that they are analyzed independently of spatial luminance contrast.

3. General properties of the model

Since this model is a first approximation whose parameters and structural details are quite open to improvement, it is worth giving a brief overview of its essential structure. The heart of the model is a set of *feature sensors*, which perform different measurements upon the input image. The set of measurements taken from a particular image make up a *feature vector*. The sensors are perturbed by noise, so that from presentation to presentation, the same image will give rise to somewhat different feature vectors. Following each trial of an experiment, the observer processes the feature vector in an optimal way to arrive at a psychophysical decision.

The model thus divides naturally into two parts: generating the feature vector, and processing the feature vector.

4. Generating the Feature Vector

Each feature sensor is defined by a *spatial weighting function*, which is a model or template of the feature to be sensed. The measurement is performed by cross-correlating the contrast image and the weighting function. The features used in the model are typified by the pattern in Fig.1(top). It is the product of a 2D sinusoid and a 2D Gaussian. Its various parameters, which may differ from sensor to sensor, are: *spatial frequency* (the frequency orthogonal to the bars), the *orientation* (the angle of the orthogonal to the bars), the *width* (the size at half height of the Gaussian defined orthogonal to the bars), the *height* (the comparable measure parallel to the bars), and the *phase* of the sinusoid (defined relative to the center of the Gaussian). Finally, each function is located at a particular point in the visual field. Of these seven parameters, we shall see that two can be defined in terms of the others, leaving five feature dimensions that may vary from sensor to sensor.

This general form of weighting function closely resembles the receptive field profiles of simple cells, the most numerous class of visual neurons in the striate cortex of cat and monkey [3,13]. Most cortical cells respond only over a modest region of space, a modest band of spatial frequencies, and a modest range of orientations [3,4,15,24]. Psychophysical data are also consistent with a moderate selectivity in space, spatial frequency and orientation [1,29,32]. The Fourier transform of the feature is shown in Fig.1(bottom). where it can be seen that the sensor also responds only over a small band of spatial frequencies and orientations.

This sort of function is often named after GABOR, who showed that (in the 1D case) it minimizes the width in both space and frequency [8]. DAUGMAN has noted the virtues of the 2D version of the Gabor function [2].

4.1. Width and Height

The data of DEVALOIS *et al.* [3] indicate simple cortical cell receptive fields tend to be slightly taller than they are wide, but including this subtlety did not seem worth the extra computational effort. Accordingly, I have equated height and width, so the supports of both pattern and transform are circular. This also allows us to specify the width of a pattern, or of its transform, by a single number: the diameter at half height. With this amendment, the weighting function for the sensor can be written

$$w(x,y) = e^{-4\ln 2(x^2 + y^2)/w^2} \cos[2\pi f(x\cos\theta + y\sin\theta) + \varphi] \qquad (1)$$

Where f is the spatial frequency, θ is the orientation, w is the width, and φ is the phase.

The width of a feature and the bandwidth of its transform are inversely related

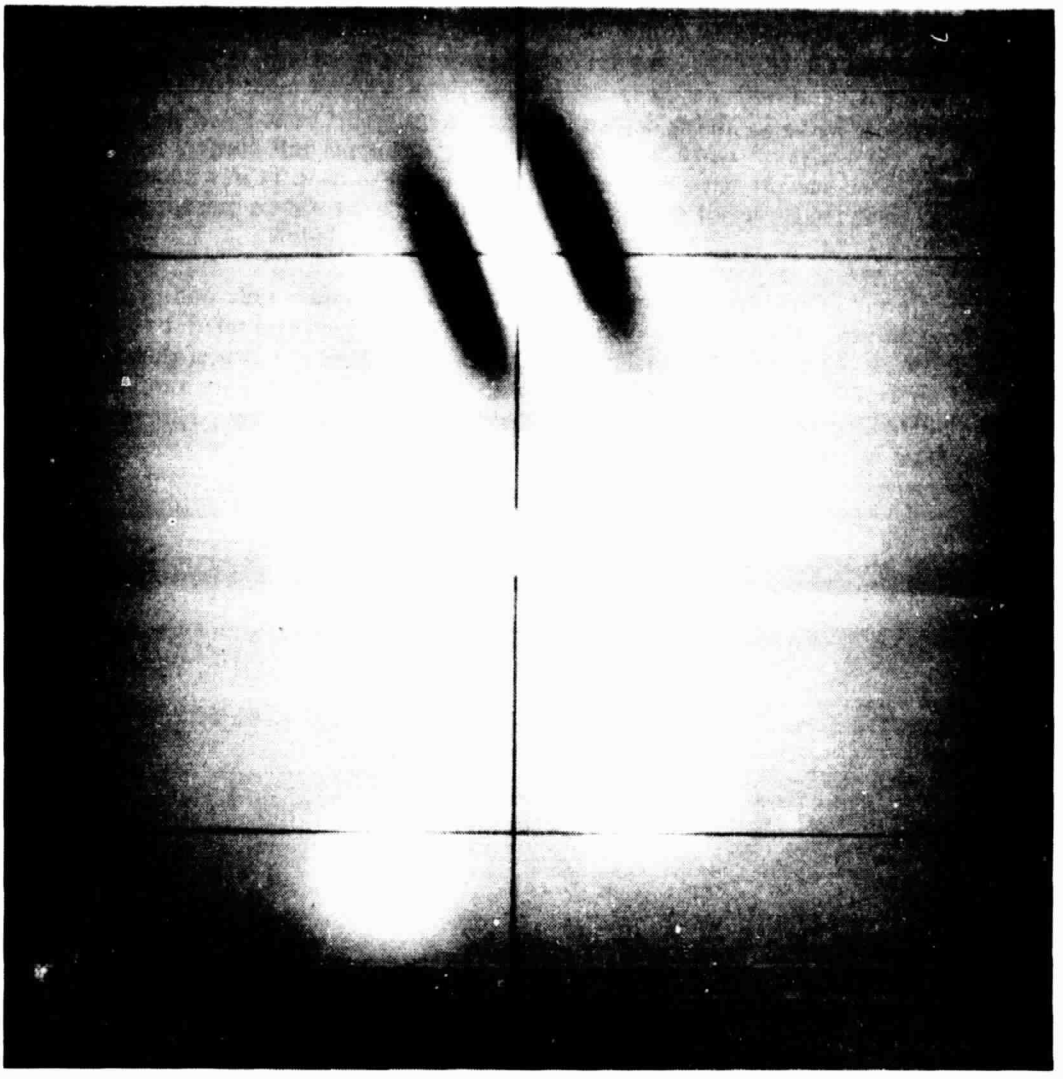

Fig.1 Top: an example of a spatial feature consisting of the product of a sinusoid and a Gaussian. The two axes are horizontal and vertical visual field position. Width of the feature is 1.324 cycles of the sinusoid. The orientation is 22.5 degrees, and the phase is 0. Bottom: the Fourier transform of the feature. The axes are horizontal and vertical spatial frequency. Top and bottom figures are to scale when 1 degree on the top scale equals 1 cycle/degree on the bottom scale.

(bandwidth = $4\ln 2/(\pi \text{width})$). I have set the width of each sensor to about 1.324 cycles of the sinusoid. This means that sensor bandwidth is proportional to sensor spatial frequency (bandwidth = 2/3 frequency), or, in logarithmic terms, that each sensor has a bandwidth of one octave. It also means that a low frequency sensor will be large, and a high frequency sensor will be small.

Spatial frequency bandwidths for simple striate cells rise approximately in proportion to frequency. Bandwidths appear to range from one half to 2.5 octaves, but one octave bandwidths comprise a large fraction of those recorded [3,15]. Psychophysical data do not lead so directly to estimates of bandwidth, but they are consistent with a proportionality between frequency and bandwidth, and with log bandwidths of about one octave [29,30,31].

4.2. Frequency

Since each sensor covers only one octave of frequency, we need a number of sensors to cover the full range of frequency which the human can sense (about 0 to 60 cycles/degree). The number required will also depend upon the density of the sensors in the frequency domain. This cannot be

less than about one sensor per bandwidth, or information in the image will be lost, and sensitivity will show dips between sensors [29]. In the model I have tried to include as few sensors as possible, so just one sensor per bandwidth has been used. This yields a set of eight basic sensors, whose frequencies range from 0.25 to 32 cycles/degree in octave steps. This basic set is shown in Fig.2. The largest sensor, a small part of which is barely visible in the upper left corner, is 128 times larger than the smallest. The sensors have been arranged in a spiral so as to fit in a square, but should all be considered to lie at the center of the visual field. This basic set will be present only at this point; the sensors at other locations will be generated by a rule described below.

The transforms of this basic set are shown in Fig.3 For clarity, only one of the two parts of each spectrum is shown. The positions of the spectra are correctly placed relative to the corresponding features in Fig.2. Note that the smallest feature corresponds to the largest transform. For clarity I have shown features with different orientations. If all the features were similarly oriented, all the transforms would lie at a single angle, and would overlap considerably. This illustrates that all of frequency space is covered.

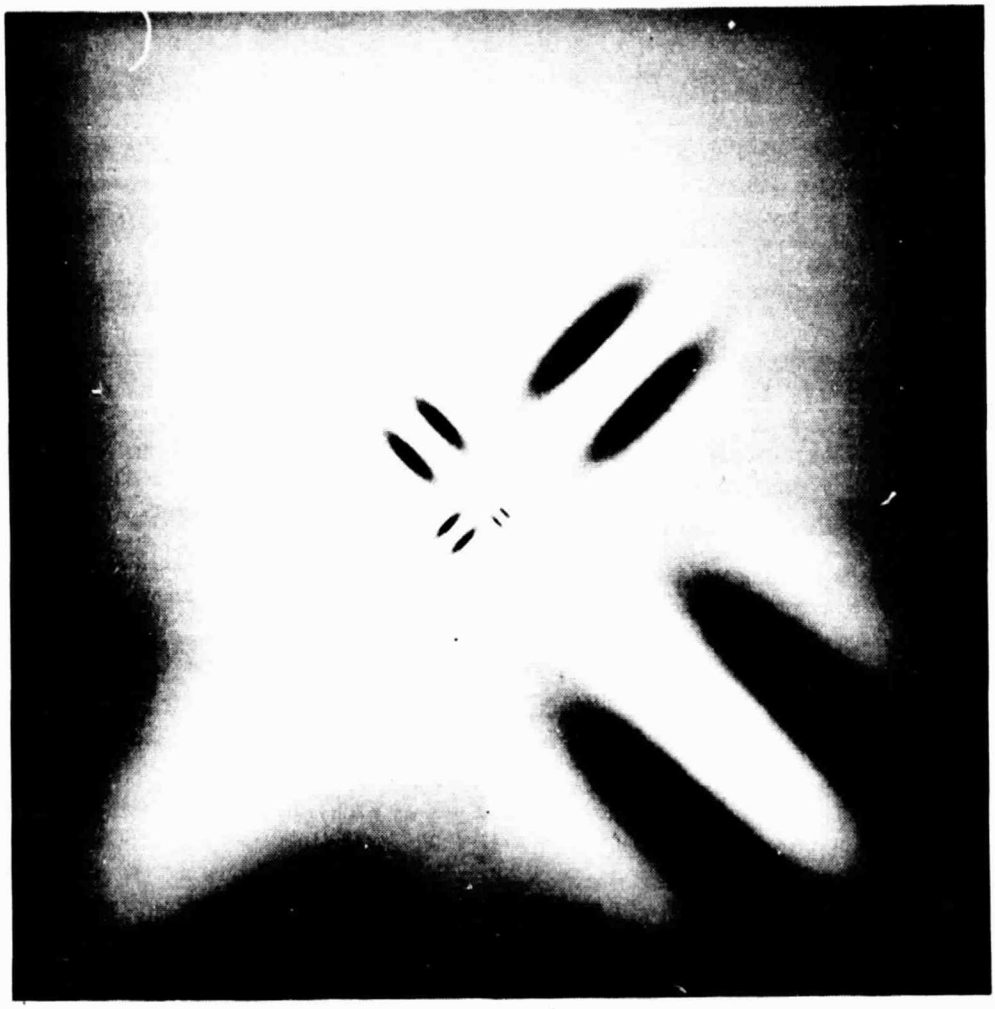

Fig.2 The basic set of sensors. The sensor frequencies range from 0.25 to 32 cycles/degree in octave steps. Each sensor has a bandwidth of one octave (width = 1.324 cycles).

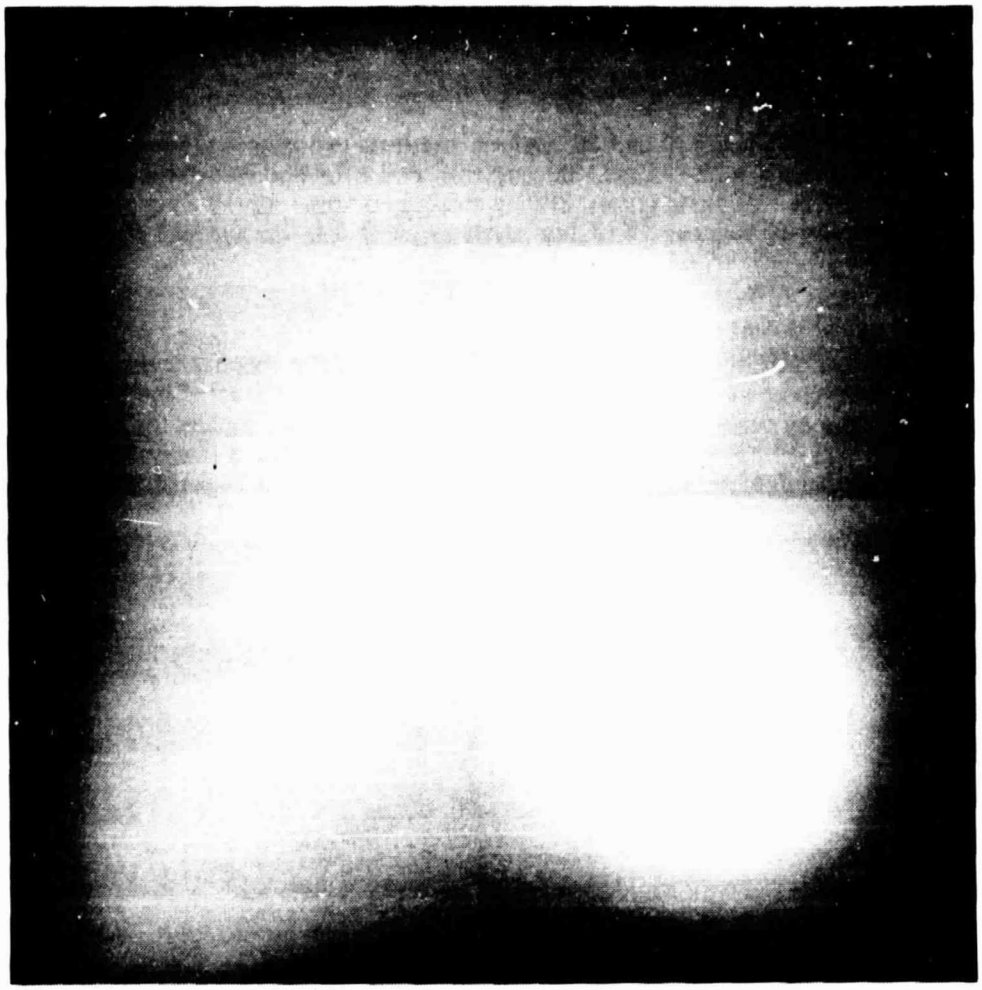

Fig.3 Fourier transforms of the set of basic sensors. Of the two Gaussians symmetrically placed about the origin, only one is shown. Axes are horizontal and vertical spatial frequency.

4.3. Spatial Sampling

To preserve the information in the image the 1D spatial density of the sensors of a particular frequency must be at least twice the frequency extent of the sensor. This extent is not well defined for a Gaussian spectrum, but if we used the bandwidth, the required sample frequency would be 4/3 the sensor frequency. I have used the more conservative factor of 1.8. Thus for the basic set of eight frequencies at the fovea, sampling intervals range from 0.0174 degrees for the highest frequency to 2.22 degrees for the lowest.

4.4. Orientation

The orientation bandwidth of a sensor (as measured with the spatial frequency to which the sensor is tuned) is completely determined by the spatial frequency bandwidth. For our one octave spatial frequency bandwidth, the orientation bandwidth is about 38 degrees. This can be seen in Fig.3, where the orientation bandwidth will be given by the range of angles that intersect the spectrum at half-height or above. This 38 degree figure agrees well with the 40 degrees reported by DEVALOIS *et al.* [4] as the median orientation bandwidth for simple cells in monkey striate cortex. There is a great deal of psychophysical data related to orientation tuning, but as yet there has been no convincing route from the data to an estimate of the orientation bandwidths of underlying sensors.

In specifying sensor density in the orientation domain I have again followed the rule of about one sensor per bandwidth. This yields the set of five orientations illustrated in Fig.4. The transforms of these features are shown in Fig.5. Successive transforms are overlaid, and each is truncated at about one width, illustrating the degree of sensor overlap in the domain of orientation.

4.5. Phase

I include sensors of two phases, 0 and 90 degrees. At least two phases are required to preserve the odd and even information in the image. Beyond this, POLLEN and RONNER [19] have found that adjacent simple cells in cat striate cortex usually share a common spatial frequency and orientation, but differ in phase by 90 degrees. Existing psychophysical data do not lead directly to any strong hypothesis regarding sensor phase.

4.6. Spatial Anisoplanatism

Perhaps the greatest difficulty in 2D modeling of human spatial vision are the large and complex variations in spatial processing that occur with distance from the center of the visual field. In general, contrast sensitivity declines with eccentricity, and does so more rapidly for high spatial frequencies than for low [11,20,27]. However it has recently been suggested that spatial processing in all portions of the visual field is identical save for a change in the 2D spatial scale [5,11,22]. This

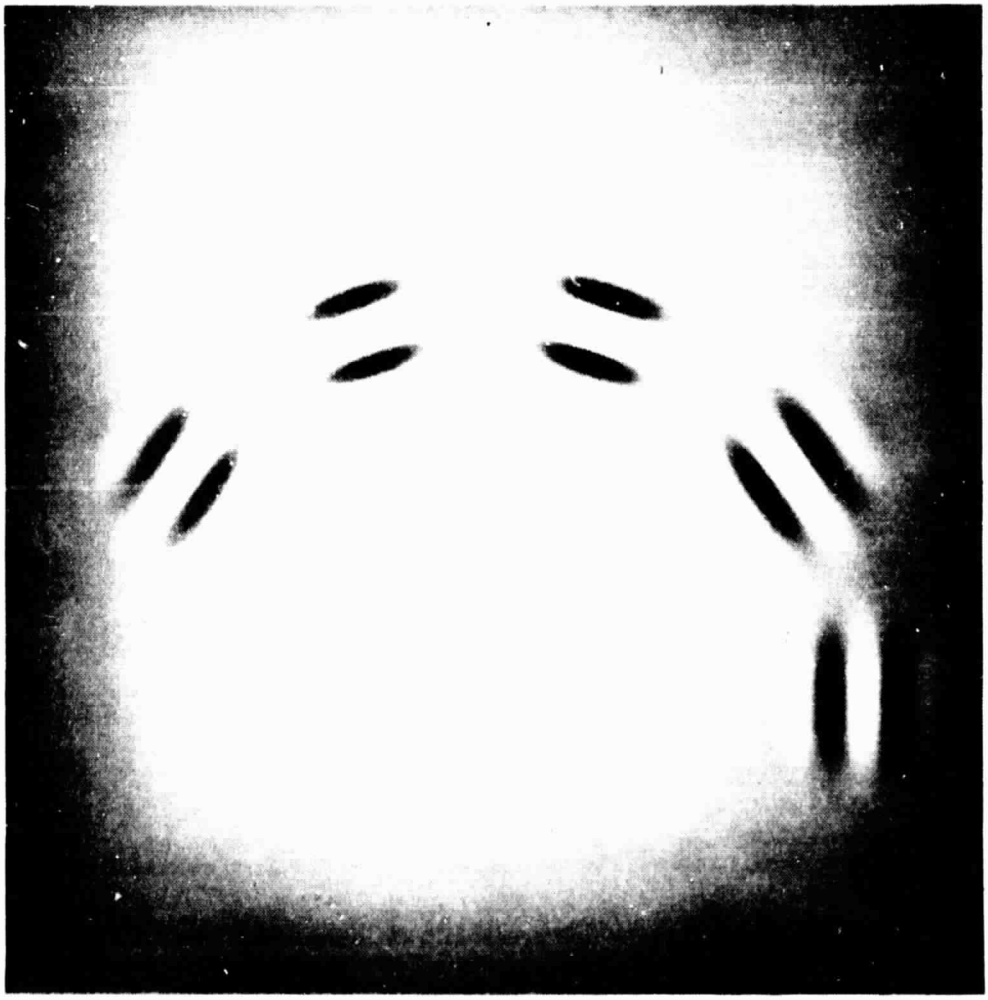

Fig.4 The five sensor orientations. They range from 0 to 144 degrees in steps of 36 degrees.

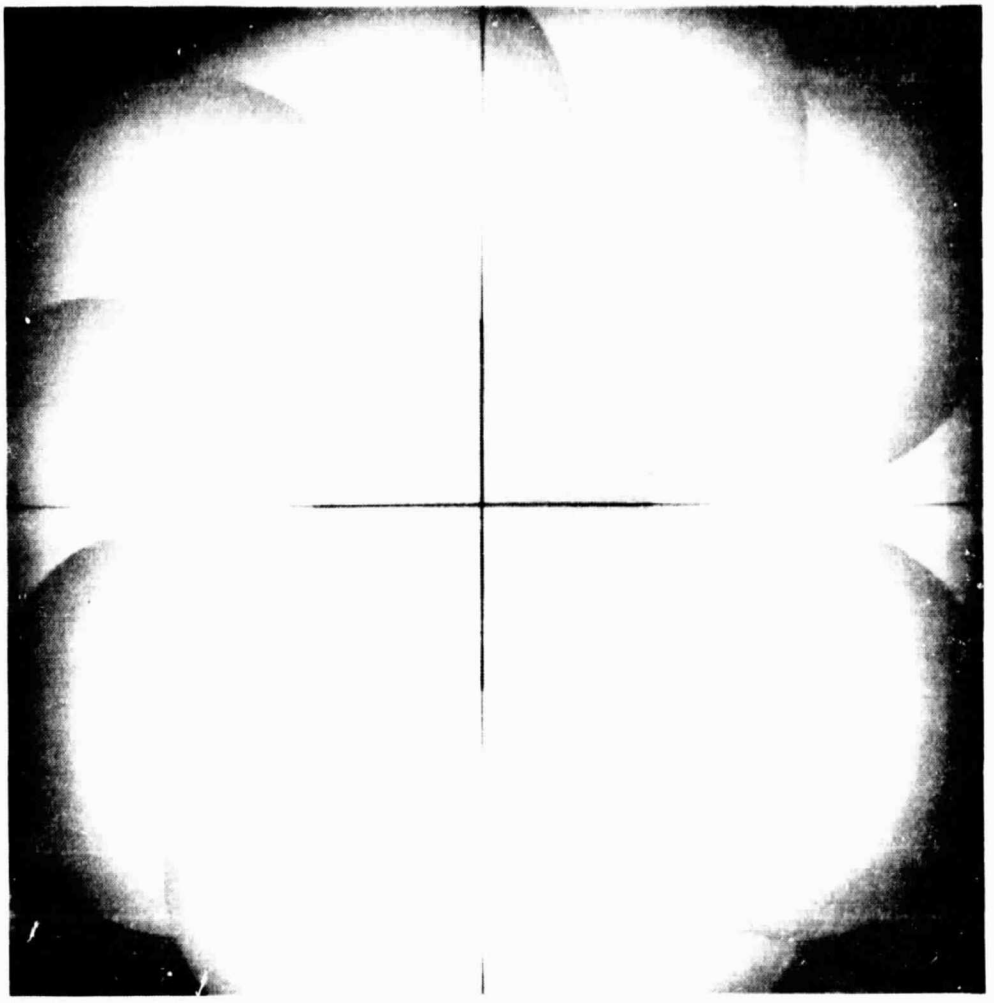

Fig.5 Transforms of the five sensor orientations. Each Gaussian has been truncated at one width for clarity. The orientation bandwidth of each sensor is about 38 degrees.

would be the case if ganglion cells changed in size with eccentricity, but all subsequent processing were homogeneous with respect to eccentricity. Though the evidence for this hypothesis is not overwhelming, it is too great a simplification to be resisted. Accordingly I have taken the spatial scale s to be given by the *scaling function*

$$s = 1 + ke \tag{2}$$

where k is a constant and e is eccentricity in degrees. For my own eyes, k is about 0.4. This agrees reasonably well with the estimates of other authors, and with the changes in ganglion cell density and cortical magnification factor with eccentricity [5,21]. It should be emphasized that this function is intended only as a first approximation.

This scaling function has been used to scale the size, frequency, and spatial density of the sensors at each eccentricity. Thus if a sensor at the fovea has a basic frequency of f width w, and 1D density d, the related sensor at eccentricity e will have frequency f/s, width ws, and density d/s. These relationships are shown in Fig.6, where I have drawn a sampling array for a basic frequency of 1 cycle/degree. Each circle in the figure is about 2/5 the width of the corresponding sensor, showing the considerable degree of spatial overlap among sensors of a common frequency. Note that as eccentricity increases, sensor frequency and spatial density are scaled by the same

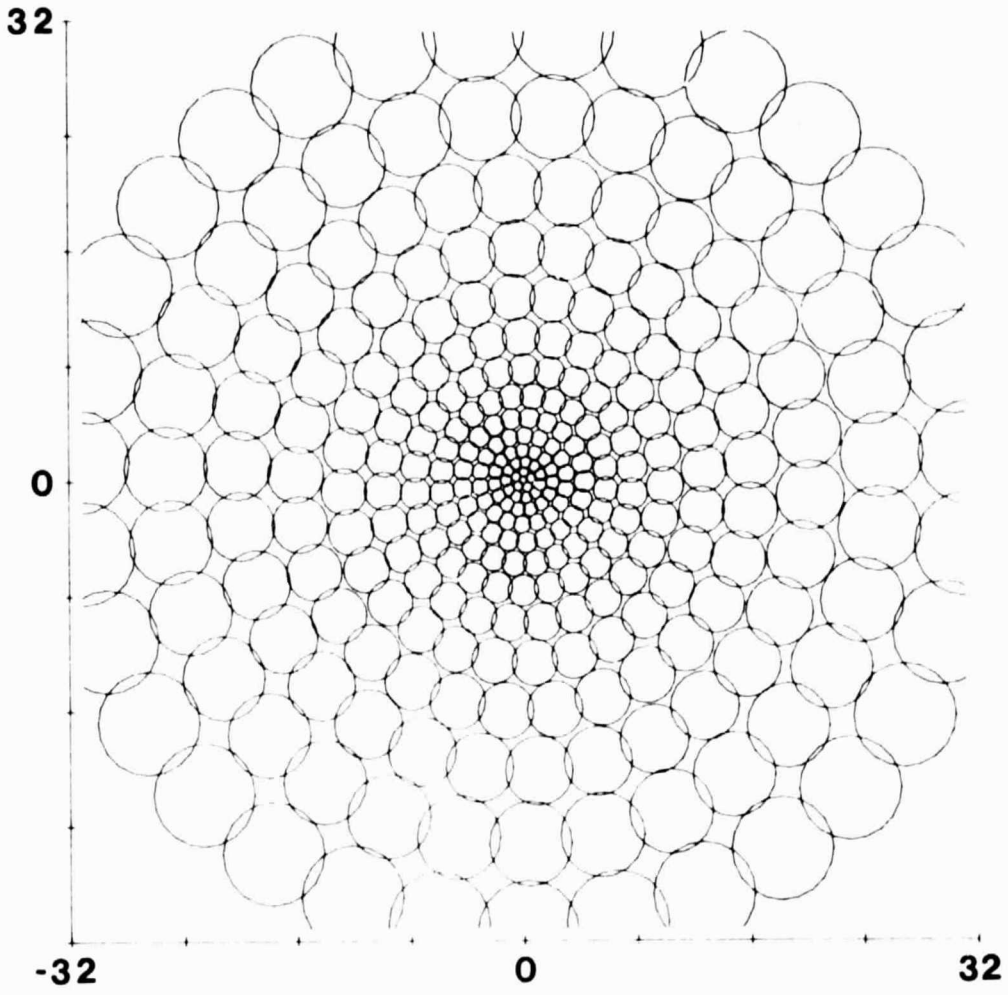

Fig.6 The sampling array for a basic frequency of 1 cycle/degree. The two axes show degrees from the fovea. Each circle is about 2/5 the width of the corresponding sensor.

amount, so that the sampling requirements continue to be met. The set of all sensors generated in this way from a given basic frequency may be considered a *family*. This is the closest we get to a "spatial frequency channel" of traditional visual theory.

The actual positions of individual sensors have been determined by finding, for each family at each eccentricity, that regular polygon centered upon the fovea whose sides are about equal in length to the difference between the present and next lesser eccentricity. This algorithm is used only as a simple way of generating sensor locations with the appropriate properties.

4.7. Sensor Gain

It is well known that contrast sensitivity depends upon spatial frequency. How do we incorporate this dependence into the model? I begin by defining *sensor gain* as the response to a matched target at unit contrast. If sensor gain at the fovea is $\chi(f)$, where f is the basic frequency, then spatial processing will be homogeneous irrespective of eccentricity only if sensor gain is adjusted to take into account the larger size of eccentric sensors. This can be accomplished by multiplying each weighting function by $\chi(f)\ 16\ ln2/(w^2\ \pi)$, where w is the actual width and f is the basic frequency of the family to which the sensor belongs. This scaling of gain would result naturally if each sensor received contributions from an equal number of receptors, and this would be likely if the receptor

density followed the same scaling function given by (2).

The actual functional form for $\chi(f)$ is not simple to estimate, since it is not given directly by any simple empirical measurement. A first approximation, indirectly estimated from contrast sensitivity to one octave bandwidth grating patches, is shown in Fig.7.

4.8. Computing the Feature Vector

The preceding assumptions permit us to compute a feature vector, z, each entry of which is the response of a single feature sensor. The many cross-correlations involved are generally more easily computed by way of well-known detours through the frequency domain. Further savings can be gained by disregarding sensors that are well outside the space or frequency support of the image. Even so, the amount of computation required to obtain the feature vector may be formidable, especially for large stimuli with high frequency content. Fortunately the sensors and all the images considered here have Fourier transforms that can be obtained analytically, so that finite transforms need not be resorted to.

5. Processing the Feature Vector

The psychological literature is filled with suggestions as to how an observer might make use of the feature vector to detect and discriminate among patterns. One might look for the largest single entry, look for the centroid of some distribution of vector entries, check whether at least one entry is larger than a criterion value, and so on. But these procedures are largely *ad hoc*, and rarely generalize much beyond the experimental context they are designed to explain. For example, looking for the largest single entry might work at threshold but is not a reasonable rule above threshold. Also, these procedures are not derived from basic principles but rather appeal to the intuition, sometimes not very forcefully.

In contrast, the statistical theory of pattern classification provides a logical basis for

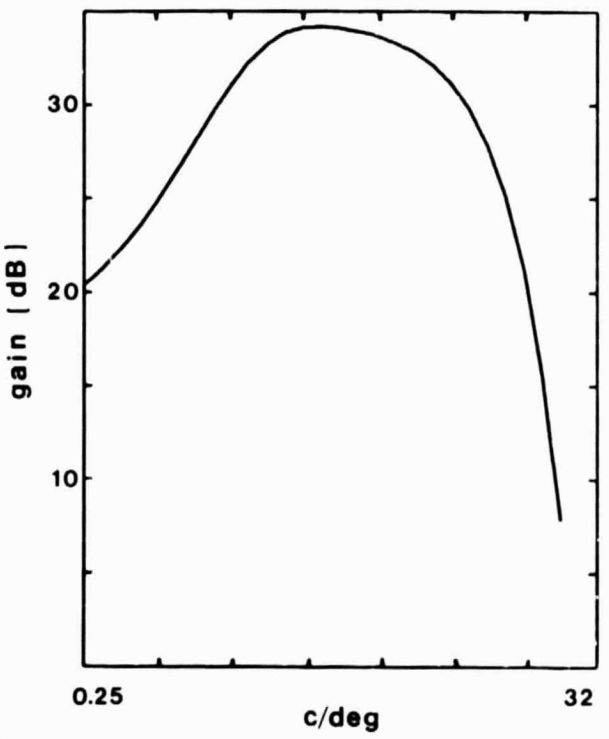

Fig.7 Sensor gain, $\chi(f)$.

constructing optimal procedures with which to analyze the feature vector [6,14]. Furthermore pattern classification theory is general, and could in principle be applied to most tasks the human observer confronts in the psychophysical lab, if not in the world at large.

5.1. The Optimal Bayesian Classifier

Let me illustrate how I apply this theory to the feature vector constructed a moment ago. First I assume that each sensor is perturbed by zero mean, unit variance Gaussian noise. The feature vector then has a multivariate normal density. I further assume that the covariance matrix of this density is the identity matrix, which means that the sensors are statistically independent.

The job of the classifier is to examine the feature vector and decide which of several possible images was in fact presented. Each of the possible images is associated with a *mean vector* $\mathbf{m}_k$. Presentation of one of the images gives rise to a feature vector $\mathbf{z}$. If we also know the prior probability of each alternative, Bayes' rule allows us to use the feature vector to calculate the posterior probability of each of the alternatives. In the cases I will consider, all alternatives have equal priors. A reasonable, and in fact optimal rule is to choose th image with highest posterior probability. Since we care only which has highest probability, it is sufficient to calculate for each alternative any quantity that is monotonic with posterior probability. Such functions on the feature vector are called *discriminants.*

In the case I have developed so far, a set of optimal discriminants are

$$d_k(z) = -(z-m_k)^T (z-m_k) \tag{3}$$

$$= -\sum_{i=1}^{n}(z_i-m_{ki})^2$$

where the superscript T denotes the transpose of a matrix. Note that if z has n elements, then it can be considered a point in n-dimensional space. Likewise for each of the mean vectors $\mathbf{m}_k$. This form of discriminant tells us to pick the alternative whose mean vector is closest to z in this space. For this reason this is often called a *minimum distance* classifier.

5.2. Prior Information

One objection to this scheme is that it assumes that the observer has perfect prior information regarding each mean vector $\mathbf{m}_k$. Since this vector is obtained through experience, and each experience is subject to variability, this assumption is unrealistic. Empirical data also clearly show that the observer is less than optimal, as though he were uncertain regarding $\mathbf{m}_k$ [7,16,25]. Fortunately, for the detection task I will consider, an approximation is available for a more realistic uncertain observer [18]. In the case of discrimination, we must be content for the moment to examine the behavior of the ideal. This is unfortunate, since to the extent that the sensors preserve the information in the image, discrimination performance tells us more about the orthogonality of the alternative images than about the properties of the sensors.

5.3. Detection

The ability of an observer to detect a pattern is often measured by a two-interval forced-choice (2IFC) method, in which the observer must judge which of two time intervals contained a signal. This may be viewed as a discrimination between two images: the test image and a null image. From a number of trials we can determine the proportion correct, or d', defined here as $\sqrt{2}$ times the normal deviate of the proportion correct. Since there are only two alternatives (image in first interval or in second interval), we can difference the two discriminants and use the sign of the result to choose an alternative. This quantity is normally distributed with a mean equal to the squared length of the mean vector, and a variance of twice the squared length. From this it is easy to show that for any given pattern,

$$d' = \sqrt{\mathbf{m}^T \mathbf{m}} \tag{4}$$

$$= \left[\sum^{n} m_i^2 \right]^{1/2}$$

As noted above, this prediction assumes an observer with perfect prior information. For an

uncertain observer, performance can be approximated by specifying that for all stimuli at some fixed d',

$$1 = \left[\sum^{n} m_i^{\beta} \right]^{1/\beta}$$

where β is a constant of about 3.5 [18]. This approximation has been used in the following predictions of detection performance.

The model gives a good account of sensitivity to a wide variety of different spatial targets. For example, the model does a reasonable job of predicting the contrast sensitivity function at various eccentricities, and the effects of size of a grating pattern upon sensitivity. Among the earliest experiments to suggest that visual mechanisms were selective for spatial frequency were those which measured sensitivity to mixtures of two frequencies [9,23]. Data from a more recent version of this experiment are shown in Fig.8 [29]. Each point shows the sensitivity to a mixture of two frequencies, relative to the sensitivity to either frequency alone. If the two frequencies add linearly, we get a ratio of two, if they do not add at all, we get a ratio of 1. The data show a decline in summation as the two frequencies move farther apart. The prediction of the model is shown by the upper curve. It does not agree precisely with the data, but given that most sources of error in the data will tend to lower the ratio it does not do a bad job. I have also shown the prediction in the case of 0.8 octave bandwidth sensors. The fit is better, and this amendment may have to be resorted to in the future.

5.4. Discrimination

A particularly sensitive way of measuring discrimination performance is by a two-by-two forced choice (2X2FC) method [17,26,31]. As in the 2IFC method, on each trial the observer is presented with an image in just one of two time intervals. But in this case the image is selected at random from a set of two. The observer must choose both the interval containing the image, and which

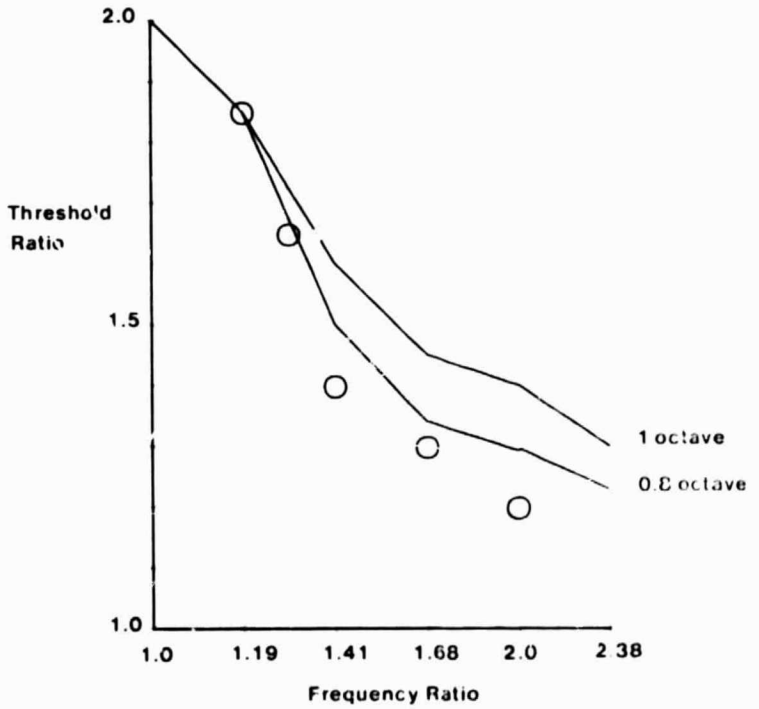

Fig.8 Summation between two different spatial frequencies. Predictions of the model with 1.0 octave and 0.8 octave bandwidths are shown. Data are from [29].

image it was. The advantage of this technique is that it measures detection and discrimination concurrently, so the two sorts of performance can be compared.

In the case of 2X2FC the observer must choose among four alternatives, so we must determine four discriminants. If we call the feature vector from the first interval $\mathbf{z}_1$ and that from the second interval $\mathbf{z}_2$, and the mean feature vectors for the two images **a** and **b**, then these discriminants are

$$d_{1a} = \mathbf{z}_1^T \mathbf{a} \qquad d_{2a} = \mathbf{z}_2^T \mathbf{a} \tag{5}$$

$$d_{1b} = \mathbf{z}_1^T \mathbf{b} \qquad d_{2b} = \mathbf{z}_2^T \mathbf{b}$$

where d_{2b} for example is the discriminant for image b in the second interval. Here we cannot reduce these to a single function of **z** as was possible in the case of 2IFC. Furthermore, although each discriminant is again normally distributed, the four are not independent, so we cannot easily derive the probability with which any one is the largest and must resort to Monte-Carlo techniques.

5.4.1. Spatial Frequency

Since our model observer discriminates among images on the basis of certain features, it is interesting to ask how well the human observer discriminates between two images that are members of the feature set. Fig.9 shows some data collected with a 2X2FC method in which the two images to be discriminated were Gaussian-windowed sinusoids, like the features of the model [31]. The ordinate plots the ratio of d' for discrimination and detection, which is a measure of discrimination performance relative to detection. This method of plotting 2X2FC data is due to THOMAS, who has done much of the pioneering work in this area. The ratio rises rapidly as the two frequencies are moved farther apart, so that when they differ by about an octave, we discriminate between them as well as we detect either one. The model's performance, shown by the solid line, is better than the human observer. Introduction of uncertainty into the model would reduce this discrepancy. Notice

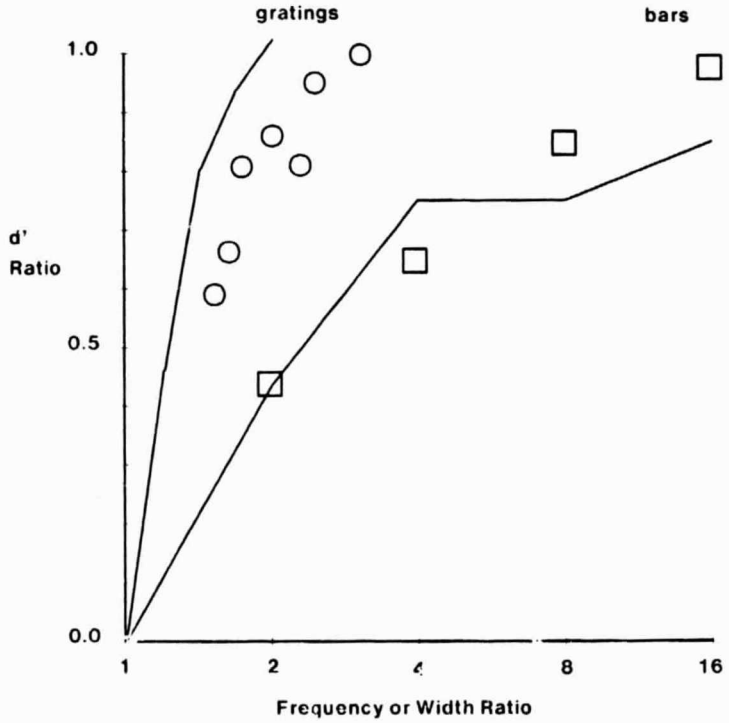

Fig.9 Discrimination between grating patches of different frequencies (circles) and Gaussian bars of different widths (squares). Curves are predictions of the model. Data are from [31] and [28].

that, as a rule of thumb, the model predicts a ratio of one when the two patterns differ by about one bandwidth.

5.4.2. Bar Width

What if we repeat the experiment with images which do not resemble the sensor features? The square symbols in Fig.9 are comparable data for discriminations between Gaussian bars: patterns with a Gaussian profile in both horizontal and vertical dimensions and with height always at least twice the width [28]. Discrimination performance is vastly poorer, so that a difference of about 4 octaves is required before discrimination is as good as detection. This poorer performance is very nicely described by the model, whose predictions are shown by the rightmost line.

5.4.3. Orientation

The last set of data I will show you were collected by THOMAS and GILLE [26]. They used a 2X2FC method to discriminate between gratings of different orientations. The gratings were large, extending 15 cycles in each direction. Their data are shown in Fig.10. Discrimination improves rapidly as the difference in orientation increases, and is almost as good as detection when the difference is 10 degrees. The model's simulated data are shown by the leftmost solid line. The agreement is quite good; perhaps too good considering uncertainty has not been included.

On the basis of a 1D model, THOMAS and GILLE estimated mechanism orientation bandwidths of 10.5 to 20.5 degrees, depending on observer. These are 2 to 4 times narrower than the sensor bandwidths used in the prediction in Fig.10. These authors acknowledged that their estimates were much narrower than bandwidths of cortical cells, and speculated that "psychophysically defined channels represent physiological mechanisms other than single cortical cells".

How can the present model with 38 degree bandwidths perform as well as the THOMAS and

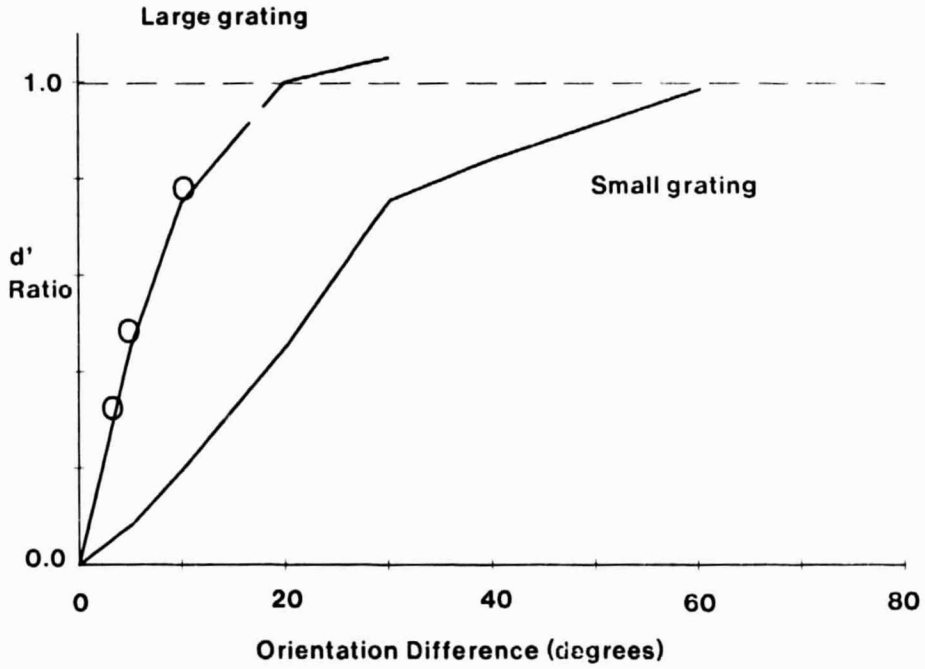

Fig.10 Discrimination between large gratings of different orientations (circles). Data are from [26]. Curves show the predictions of the model for large and small gratings.

GILLE model with bandwidths about 3 times narrower? The most likely answer is that the present model is 2D, and can make intelligent use of the pattern of activity across a large number of spatially distributed sensors. When the stimulus is confined to a small area, performance deteriorates markedly, as shown by the rightmost curve in Fig.10. (though much of this is due to the broadened orientation bandwidth of the stimulus itself).

These observations illustrate an important point: sensor properties cannot be estimated without an explicit, computable, 2D model of the sensors and of their distribution over the visual field. The model must also allow the observer to make intelligent use of the sensor outputs. This argument applies to any effort to derive sensor parameters from psychophysical data, that is, to almost all psychophysical research.

Acknowledgements

I thank Albert Ahumada for useful discussions and help in the construction of Fig.s 1-5.

References

1. C. Blakemore, F. W. Campbell *J. Physiol.*, **203**, 237-260 (1969).
2. J. D. Daugman *Vision Res.* **20**, 847-856 (1980).
3. R. L. De Valois, D. G. Albrecht, L. G. Thorell *Vision Res.* **22**, 545-559 (1982).
4. R. L. De Valois, E. W. Yund, N. Helper *Vision Res.* **22**, 531-544 (1982).
5. N. Drasdo *Nature* **266**, 554-556 (1977).
6. R. O. Duda, P. E. Hart *Pattern classification and scene analysis.* (John Wiley, New York, 1973)
7. J. M. Foley, G. E. Legge *Vision Res.* **21**, 1041-1053 (1982).
8. D. Gabor *J. IEE London*, **93**(III), 429-457 (1946).
9. N. Graham, J. Nachmias *Vision Res.*, **11**, 251-259 (1971).
10. D. M. Green, J. A. Swets *Signal Detection Theory and Psychophysics.* (Wiley, New York, 1966).
11. J. J. Koenderink, M. A. Bouman, A. E. Bueno de Mesquita, S. Slappendel *J. opt. Soc. Amer.* **68**, 845-849 (1978).
12. J. J. Koenderink, M. A. Bouman, A. E. Bueno de Mesquita, S. Slappendel *J. opt. Soc. Amer.* **68**, 854-860 (1978).
13. S. Marcelja *J. Opt. Soc. Am.* **70**, 1297-1300 (1980).
14. J. G. Moik *Digital processing of remotely sensed images.* (NASA SP-431, Washington, 1980).
15. J. A. Movshon, I. D. Thompson, D. J. Tolhurst *J. Physiol.*, **283**, 53-77 (1978).
16. J. Nachmias, E. C. Kocher *J. Opt. Soc. Am.*, **60**, 382-389 (1970).
17. J. Nachmias, A. Weber *Vision Res.*, **15**, 217-233 (1975).
18. D. G. Pelli *Effects of visual noise.* (Ph.D. Thesis, University of Cambridge, Cambridge, England, 1981).
19. D. A. Pollen, S. F. Ronner *Science* **212**, 1409-1411 (1981).
20. J. G. Robson, N. Graham *Vision Res.* **21**, 409-418 (1981).
21. J. Rovamo, V. Virsu *Exp. Brain Res.* **37**, 495-510 (1979).
22. J. Rovamo, V. Virsu, R. Nasaren *Nature* **271**, 54-56 (1978).
23. M. B. Sachs, J. Nachmias, J. G. Robson *J. opt. Soc. Am.* **61**, 1176-1186 (1971).
24. P. H. Schiller, B. L. Finlay, S. F. Volman *J. Neurophysiol.* **39**, 1334-1351 (1976).
25. W. P. Tanner Jr., J. A. Swets *Transactions of the IRE*, **PGIT-4**, 213-221 (1954).
26. J. P. Thomas, J. Gille *J. Opt. Soc. Am.* **69**, 652-660 (1979).
27. A. J. van Doorn, J. J. Koenderink, M. A. Bouman *Kybernetik* **10**, 223-230 (1972).

28. A. B. Watson *Optical Society of America, Technical Digest of Topical Meeting on Recent Advances in Vision,* SA4 (1980).

29. A. B. Watson *Vision Res.* **22**, 17-25 (1982).

30. A. B. Watson, H. B. Barlow, J. G. Robson *Invest. Ophthlmol. and Visual Science* Suppl. to **20**, 178 (1981).

31. A. B. Watson, J. G. Robson *Vision Res.*, **21**, 1115-1122 (1981).

32. D. W. Williams, H. R. Wilson, J. D. Cowan *J. opt. Soc. Amer.* **72**, 878-887 (1982).