Perception, 1996, volume 25, pages 335–342

Regularity vs genericity in the perception of collinearity

Jacob Feldman
Department of Psychology, Center for Cognitive Science, Rutgers University, New Brunswick, NJ 08903, USA
Received 17 January 1995, in revised form 7 November 1995

Abstract. The perception of collinearity is investigated, with the focus on the minimal case of three dots. As suggested previously, from the standpoint of probabilistic inference, the observer must classify each dot triplet as having arisen either from a one-dimensional curvilinear process or from a two-dimensional patch. The normative distributions of triplets arising from these two classes are unavailable to the observer, and are in fact somewhat counterintuitive. Hence in order to classify triplets, the observer invents distributions for each of the two opposed types, 'regular' (collinear) triplets and 'generic' (ie not regular) triplets. The collinear prototype is centered at 0° (ie perfectly straight), whereas the generic prototype, contrary to the normative statistics, is centered at 120° away from straight—apparently because this is the point most distant in triplet space from straight and thus creates the maximum possible contrast between the two prototypes. By default, these two processes are assumed to be equiprobable in the environment. An experiment designed to investigate how subjects' judgments are affected by conspicuous environmental deviations from this assumption is reported. The results suggest that observers react by elevating or depressing the expected probability of the generic prototype relative to the regular one, leaving the prototype structure otherwise intact.

1 Collinearity detection as probabilistic inference

The extraction of curvilinear structure from a field of localized visual items (eg dots) has been modeled in a number of different ways: as the output of orientation-tuned local operators (Glass 1969; Caelli and Julesz 1978; Prazdny 1984); by reflexive reasoning principles (eg Gestalt rules); and by computational methods implementing these and other concepts (Stevens 1978; Zucker 1985; Stevens and Brookes 1987; Parent and Zucker 1989). I previously suggested (Feldman 1993) that the process can also be regarded as a kind of probabilistic inference. This idea refers to the logic of the collinearity inference, not of the process by which it is carried out, and thus does not necessarily contradict any of the above models per se. Nevertheless it does provide some novel predictions concerning the exact magnitude of collinearity judgments (see Feldman 1993). In this paper I report an experiment corroborating one of these predictions.

Consider an observer presented with some set of dots, which he or she may or may not regard as falling in a curvilinear sequence (figure 1). The observer in effect has to decide between two alternative generating processes as competing 'explanations' for the observed dots (cf Leyton 1984, 1992). Each alternative hypothesis is a region inside of which dots are generated in a probabilistic fashion (eg what statisticians call a Poisson process). But is this generating process a one-parameter curve or a two-dimensional patch? The former hypothesis is a regular pattern that imposes some structure on the dots, while the latter hypothesis has no internal structure at all—a random null hypothesis. The three-dot case represents the minimal decision (because all pairs of points are equivalent up to translation, rotation, and scaling). Here the observer is presented with three dots (figure 2) and must decide, in effect, whether the configuration is collinear 'enough' to justify the inference of an underlying curve.

This problem seems at first glance like a well-defined probabilistic decision, and it makes sense to begin by approaching it from a normative point of view. For simplicity, we consider only one variable, the angular deviation of the dot triplet from perfect

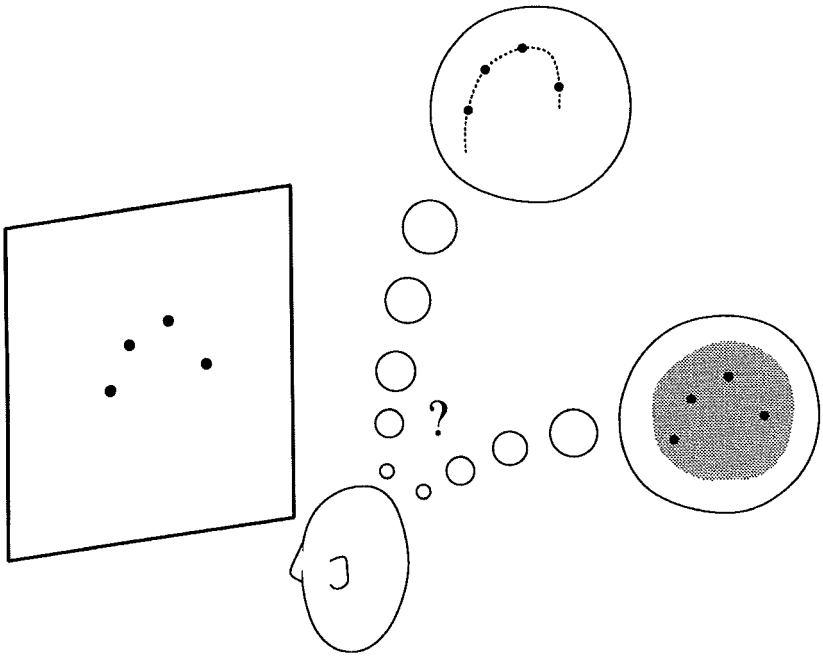

Figure 1. Shown a field of dots, the observer is faced with a choice: did the dots originate from a process generating a one-dimensional curve (top), or from a process generating a two-dimensional patch (bottom).

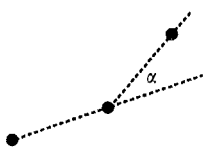

Figure 2. A typical stimulus, with the angle α marked.

collinearity, denoted α (so $\alpha = 0°$ means a collinear triplet, while $\alpha = 120°$ means an equilateral triangle when interdot distances are equal). There are two generating distributions. The first is the curve distribution, (α); this measures the expected angle when three dots are sampled 'at random' from a smooth curve which itself was chosen 'at random'. The second is the cluster distribution $\mathcal{K}(\alpha)$; this measures the expected angle of three dots that were generated at random in the plane. Because here the three dots are inherently unordered, we must pick an order in which to measure angle, say, the most nearly collinear path.

Assuming the two generating distributions $\mathcal{C}(\alpha)$ and $\mathcal{K}(\alpha)$ occur with equal probability in the environment, the posterior probability that the three observed dots actually originated from the curve process is simply

$$p(\mathcal{C}|\alpha) = \frac{\mathcal{C}(\alpha)}{\mathcal{C}(\alpha) + \mathcal{K}(\alpha)} \tag{1}$$

by Bayes's theorem.

The cluster distribution $\mathcal{K}(\alpha)$ has been extensively investigated by the statisticians Kendall and Kendall (1980) for the case where three dots are generated from a circular Gaussian in the plane (see also Broadbent 1980). The distribution is bounded below by 0° and above by 120°. It is somewhat counterintuitive in form; it peaks at 90°, and precisely 120° triplets are much more rare than precisely 0° triplets (see figure 3a).

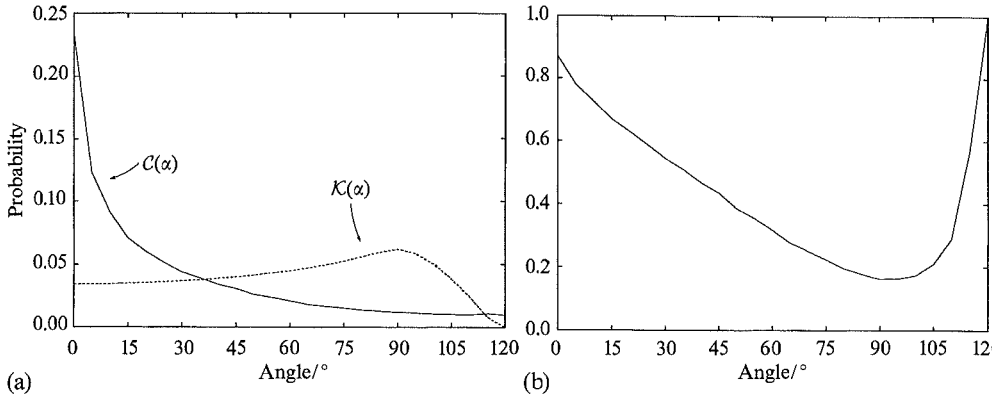

Figure 3. (a) 'Normative' distributions of triplets of dots sampled from a smooth curve [$C(\alpha)$] and randomly generated in the plane [$K(\alpha)$]. (b) The resulting posterior distribution $p(C|\alpha)$.

The curve distribution $C(\alpha)$ is more difficult to characterize in a strictly normative way, because of the difficulty in defining a 'random' sample from a 'random' curve. The general idea is to choose a smooth curve, and then pick three points at random intervals of arc length. The resulting distribution thus depends to a large extent on the contingent distribution of curves in the environment, the intervals at which they tend to be sampled, and so forth.

Nevertheless, the assumption that the curves are *smooth* tells us something about the distribution of triplets: that it can be expected to be sharply peaked at 0°. To see why, consider triplets sampled at finer vs coarser scales (figure 4). Within any sufficiently small neighborhood, finer scales always lead to more nearly straight triplets; in fact this is virtually the definition of smoothness, ie that as the neighborhood gets smaller the curve is better and better approximated by a local tangent. By contrast, coarser scales do not necessarily yield less-straight triplets; this depends on the global structure of the curve, which is arbitrary. Integrating across all scales, smaller scales contribute more univocally, and hence straighter triplets tend to dominate the resulting distribution.

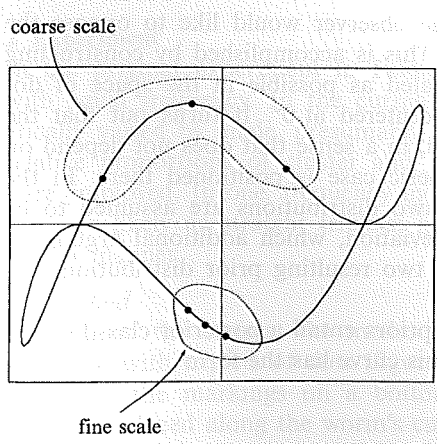

Figure 4. A smooth curve, showing how triplets sampled at a fine scale tend to produce relatively straight triplets, while triplets sampled at a coarse scale can produce triplets at arbitrary angles. The curve shown here is from the parametric family $x(t) = A\cos(t) + B\sin(t)$, $y(t) = C\cos(t)\sin^2(t)$, chosen to have a balance of high-curvature and low-curvature segments. Here $A = B = C = 1$. The same curve family is used in the simultation shown in figure 3a, with values of A, B, and C selected randomly on each trial.

A simulation was carried out in order to corroborate this argument, and to uncover the structure of the distribution in more detail. First, a parametric family of curves was defined, constructed so as to include regions of both high and low curvature. On each trial, a triplet was generated by sampling points at a randomly chosen scale from a randomly chosen member of the curve family. The results of 200 000 such trials are plotted as $\mathcal{C}(\alpha)$ in figure 3a. The curve is clearly peaked at 0°, and trails off monotonically at greater angles. Notice that large angles (eg 120°), while relatively rare, do occur.

An observer wishing to classify dot triplets normatively as having originated from either a curve process or a patch process would have to use the above distributions $\mathcal{C}(\alpha)$ and $\mathcal{K}(\alpha)$ as priors (see Ashby and Gott 1988). The resulting posterior curve $p(\mathcal{C}|\alpha)$ is shown in figure 3b. However, human observers are presumably ignorant of the formal statistics of the situation, and what classification rule they actually use is an empirical question.

I previously reported (Feldman 1993) an experiment in which subjects were shown configurations of three dots and asked (via suitably comprehensible instructions) whether the dots originated from a curve or randomly in the plane. Their judgments do not resemble the normative curve of figure 5. Instead, the 'curvilinear' judgment is maximum at 0° and tails off to a minimum at 120°.

1.1 *Regularity vs genericity*

Subjects' actual classification curve can be explained completely if one assumes that human observers construct triplet distributions by using certain heuristic principles, hinging on the dichotomous opposition of 'regular' and 'generic' forms. As argued in Feldman (1993), collinearity can be regarded as a *regularity*—a configuration whose high degree of rule-like structure serves as a clue to the presence of some distal organizing process, in this case the curvilinear trace (Lowe 1987). For the observer seeking to detect a distal curve, then, the collinear triplet is the target. The natural contrast to this regular case is the irregular or 'generic' case: three dots in general position without any special structure. Genericity—a mathematical generalization of the idea of 'typicality'— is an important concept both in differential geometry (Poston and Stewart 1978; Koenderink 1990) and in perceptual theory (Binford 1981; Bennett et al 1989; Richards and Jepson 1992).

In order to maximize inferential leverage, the observer would like to oppose the regular and generic cases as starkly as possible. This is accomplished by constructing two prior distributions that are as widely separated as possible in the space of dot triplets. The collinear distribution, naturally, is centered at 0°. It turns out that the point most distant from 0° in the triplet manifold, in a sense that does not depend on the choice of metric, is at 120°. Hence the generic case is positioned there. In the absence of any arguments to the contrary, the two distributions are assumed to be Gaussians with the same height and standard deviation, which additional arguments fix at about 53° (again, see Feldman 1993). The two resulting prior distributions are shown in figure 5a.

As with the normative priors $\mathcal{C}$ and $\mathcal{K}$, these priors entail a posterior classification curve $p(\text{'curve'}|\alpha)$, which is shown in figure 5b. This curve has the form

$$p(\text{'curve'}|\alpha) = \frac{N(0°, 53°)}{N(0°, 53°) + \lambda N(120°, 53°)}, \tag{2}$$

where $N(\mu, \sigma)$ is the Gaussian with mean μ and standard deviation σ, and λ is the ratio in the environment of the population of cluster processes to curve processes. This curve matches subjects' judgments in the triplet-classification task almost perfectly ($R^2 = 0.99$) with λ set to the default value of unity.

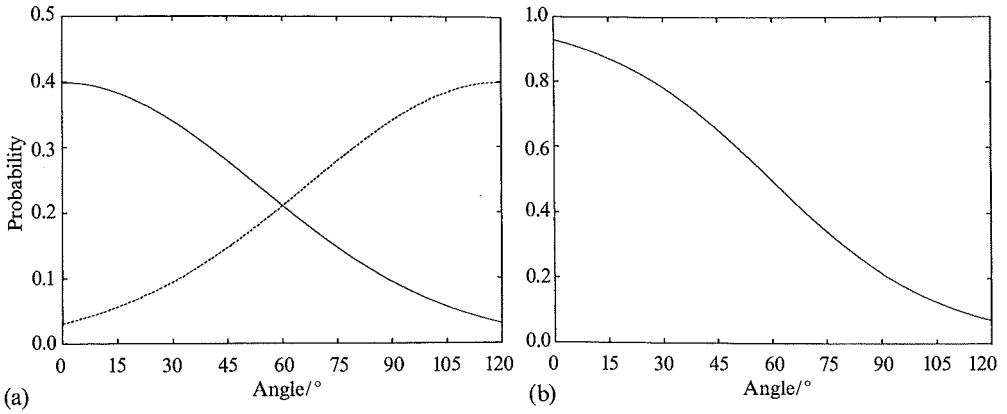

Figure 5. (a) Theoretical priors for the curve and cluster distributions, derived so as to maximize the opposition of regular and generic forms. (b) The resulting posterior classification function $p(\text{'curve'}|\alpha)$.

It should be emphasized that this interpretation scheme, while it may seem intuitively reasonable, is extra-Bayesian—in that the priors are imported via non-Bayesian arguments—and highly counternormative. Subjects apparently regard equilateral (120°) triplets as very likely outcomes of a random cluster process—indeed, the very prototype of such a process—but extremely unlikely outcomes of a curve process. But in fact such triplets are *extremely* unlikely outcomes of a cluster process, but only *mildly* unlikely outcomes of a curve process.

This account of collinearity classification gives a central role to subjects' prior assumptions about the relative probabilities of various processes and configurations in the world. Hence it is reasonable to wonder how subjects' judgments are affected by observed environmental deviations from these assumptions. For example, the empirical fit to equation (2) implies that subjects tacitly assume that cluster processes and curve processes are equally common in the world—ie that $\lambda = 1$. This is a contingent assumption, and it is easy to construct an artificial world in which it is violated. In the following experiment this possibility is pursued.

2 Experiment

In the original experiment in Feldman (1993), triplets of all angles were presented with equal frequency. In the current experiment the frequency of triplets is systematically varied by angle, presumably biasing subjects to change their beliefs about the relative proportions of cluster and curve processes in the environment. Two scripts were employed. In one, frequency increased with angle so that equilateral triplets were twice as likely as straight triplets; in the other, frequency decreased with angle in a similar manner.

2.1 *Method*

Aside from the scripts, the methodology was as in Feldman (1993). Subjects were presented with triplets of dots (figure 2). They were instructed that in some cases the dots were the markings on a bottom-dwelling species of ocean worm, which were randomly placed along the worm's curvilinear body. In other cases, however, the three dots were simply randomly placed rocks on the ocean floor. The instructions explicitly drew subjects' attention to the fact that any triplet could have either a 'worm' or 'random-rocks' interpretation: worms can curl up arbitrarily, while rocks might just 'happen' to fall nearly in a line. As illustration, subjects were shown two worms and random-rock configurations, arranged so that exactly the same two dot triplets were

given as an example of both categories. On each trial, subjects responded by choosing a number from 1 to 5 representing their confidence that the observed triplet was a worm.

2.1.1 *Subjects.*
Twenty subjects from the university community were paid for their participation. Ten were assigned randomly to script 1 and ten to script 2.

2.1.2 *Stimuli.*
In script 1, there were eight triplets at 0°, and then one additional triplet for each 15° increment away from 0° both in the positive and in the negative direction, to a maximum of sixteen triplets at both 120° and −120°. This scheme was crossed with two levels of interdot distance: equal, and unequal with a ratio chosen randomly from values between 1 and 2. This amounts to a total of 416 trials. Script 2 followed the same scheme except with frequency decreasing from sixteen at 0° to eight at 120° and −120°, again multiplied by two levels of interdot distances; this amounts to a total of 400 trials [less than script 1 because the most numerous case there occurs twice (120° and −120°), while in script 2 the most numerous case (0°) only occurs once].

Triplets were each rotated in the plane by a random angle. Dots were black circular patches 5 pixels in diameter (subtending 0.23 deg of visual angle at 35 cm distance) on a white screen with high contrast. Interdot distance ranged from 100 pixels (4.5 deg) to 300 pixels (13.5 deg). Subjects were free to take as long as they wanted to respond.

2.2 *Results*
The results for both scripts are displayed in figure 6, displaying the probability of a 'worm' response (average response collapsed across subjects) as a function of angle. Along each set of data, the best-fit curve is plotted from the family of posterior curves represented by equation (2) with σ fixed at 53.059° [the actual empirical value found in Feldman (1993)], but with λ left free to vary.

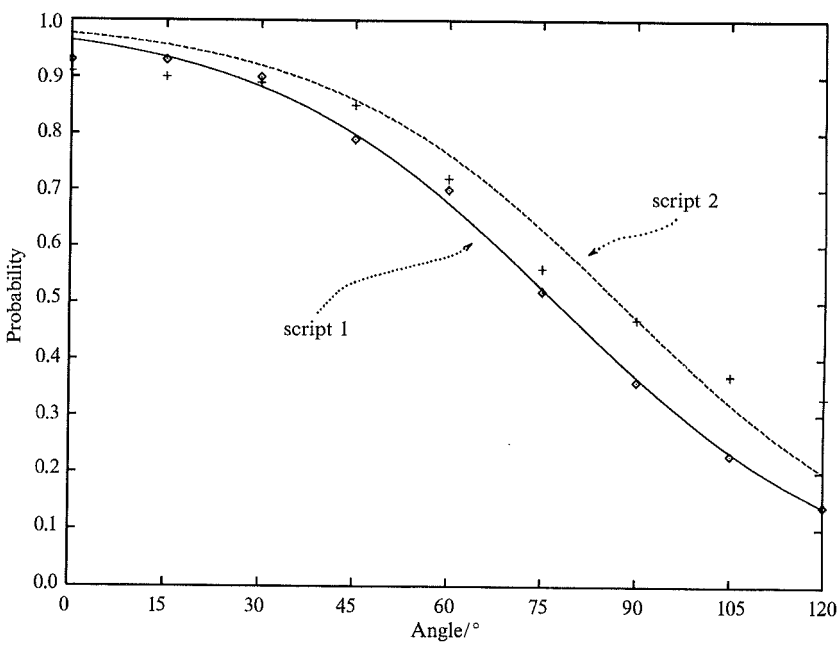

Figure 6. Results from the experiment (probability of a 'worm' response plotted as a function of angle) for both script 1 (diamonds) and script 2 (crosses). For each case, also plotted is the best-fit models from the posterior family parameterized by λ as described in the text.

This model fits both sets of data extremely well. In the rising case (script 1), the curve fit with $R^2 = 0.999$, $F_{1,7} = 12.930$, $p \ll 0.01$ with λ fit to 0.478. In the falling case (script 2), $R^2 = 0.992$, $F_{1,7} = 891.904$, $p \ll 0.01$, with λ fit to 0.311.

2.3 *Discussion*

The results corroborate the regularity-vs-genericity account of subjects' collinearity judgments, with only the exact proportions of the two canonic types being varied. Under the influence of the two scripts, subjects have simply varied their estimate of the relative frequency of curve and cluster processes in their environment. That is, they have constructed the two class densities exactly as before—counterposing generic and nongeneric categories—and then simply adjusted the height of the generic 'bin' as compared with the nongeneric bin in order to reflect better the conspicuously nonuniform environment in which they found themselves situated. Oddly, they seem to have estimated the former to be more probable in both conditions; hence, while they move bins around in reaction to the environment, they do not seem to do so in an accurate manner. They varied an existing parameter in the posterior calculation, but otherwise preserved unchanged the basic structure of the categorization observed in the original experiment.

3 Conclusions

From the experiment reported here, coupled with that in Feldman (1993), it is argued that collinearity detection can be usefully regarded as a probabilistic inference provided that observers make some decidedly extra-Bayesian assumptions. Subjects unconsciously construct two maximally contrasted prototypes: one of regularity, the other of genericity. The regular prototype—the nearly collinear triplet—signals the presence of a distal curve process. The generic prototype—the nearly equilateral triplet, because it is the most distant point from collinearity in the triplet manifold—signals the absence of any such regular process. As demonstrated above, this latter assumption is not normative. Nevertheless, it makes some sense if one regards the observer as above all a seeker of pattern and structure in the image, which constitute its only proximate clues about distal structures (Witkin and Tenenbaum 1983; Richards et al 1996). This argument suggests that it might be possible to express grouping interpretations—not just collinearity judgments—entirely in terms of the detection of regular patterns, by using some more generalized definition of 'regular' than that used in this paper. Indeed I have presented such a model (Feldman 1995b), in which Boolean decisions about the presence or absence of regularities in the image are composed in a systematic way into an overall interpretation of structure. In this model the perceived grouping interpretation turns out to be the most regular interpretation drawn from a well-defined logical language of grouping interpretations.

To understand more completely human judgments of curvilinearity in a probabilistic framework, the next step is to move beyond the minimal three-dot case to four dots and beyond (pursued in Feldman 1995a). For the observer interpreting a chain of dots, each interleaved triplet case can be thought of as a single 'trial' in a series of potentially independent trials. Any sequence of four or more dots corresponds to a sequence of such trials. Hence the judgment of curvilinearity in such cases hinges on the observer's expectations about the *joint* behavior of successive interleaved triplets. Mathematically, any interdependence between successive triplets in a chain can be expressed literally as covariance between successive angles—another type of 'regularity'. In this way, individual triplet classifications of the type investigated above are combined into a global inference of a distal curve.

Acknowledgements. This research was made possible by the Rutgers Center for Cognitive Science, New Brunswick, NJ. I am very grateful to Whitman Richards, Allan Jepson, and several anonymous reviewers for many helpful comments, and to Sheryl Maniar for collecting data from the subjects.

References

Ashby F G, Gott R E, 1988 "Decision rules in the perception and categorization of multidimensional stimuli" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **14** 33–53

Bennett B M, Hoffman D D, Prakash C, 1989 *Observer Mechanics: A Formal Theory of Perception* (London: Academic Press)

Binford T, 1981 "Inferring surfaces from images" *Artificial Intelligence* **17** 205–244

Broadbent S, 1980 "Simulating the ley hunter" *Journal of the Royal Statistical Society, Series A* **143** 109–140

Caelli T M, Julesz B, 1978 "On perceptual analyzers underlying visual texture discrimination: Part 1" *Biological Cybernetics* **28** 167–175

Feldman J, 1993 "Perceptual models of small dot clusters", in *Proceedings of the DIMACS Workshop on Partitioning Data Sets* (Providence, RI: American Mathematical Society) pp 331–357

Feldman J, 1995a "Collinearity, covariance, and regularity in perceptual groups" (submitted to *Vision Research*)

Feldman J, 1995b "Regularity-based perceptual grouping", technical report 21, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ

Glass L 1969 "Moiré effects from random dots" *Nature (London)* **223** 578–580

Kendall D G, Kendall W S, 1980 "Alignments in two-dimensional random sets of points" *Advances in Applied Probability* **12** 380–424

Koenderink J J, 1990 *Solid Shape* (Cambridge, MA: MIT Press)

Leyton M, 1984 "Perceptual organization as nested control" *Biological Cybernetics* **51** 141–153

Leyton M, 1992 *Symmetry, Causality, Mind* (Cambridge, MA; MIT Press)

Lowe D G, 1987 "Three-dimensional object recognition from single two-dimensional images" *Artificial Intelligence* **31** 355–395

Parent P, Zucker S W, 1989 "Trace inference, curvature consistency, and curve detection" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11** 823–839

Poston T, Stewart I N, 1978 *Catastrophe Theory and Its Applications* (London: Pitman)

Prazdny K, 1984 "On the perception of glass patterns" *Perception* **13** 469–478

Richards W, Jepson A, 1992 "What makes a good feature?", in *Spatial Vision in Humans and Robots* Eds L Harris, M Jenkins (Cambridge: Cambridge University Press)

Richards W, Jepson A, Feldman J, 1996 "Priors, preferences, and categorical percepts", in *Perception as Bayesian Inference* Eds D Knill, W Richards (Cambridge: Cambridge University Press)

Stevens K A, 1978 "Computation of locally parallel structure" *Biological Cybernetics* **29** 19–28

Stevens K A, Brookes A, 1987 "Detecting structure by symbolic constructions on tokens" *Computer Vision, Graphics, and Image Processing* **37** 238–260

Witkin A P, Tenenbaum J M, 1983 "On the role of structure in vision", in *Human and Machine Vision* Eds J Beck, B Hope, A Rosenfeld (New York: Academic Press) pp 481–543

Zucker S W, 1985 "Early orientation selection: Tangent fields and the dimensionality of their support" *Computer Vision, Graphics, and Image Processing* **32** 74–103

p

© 1996 a Pion publication printed in Great Britain