Contour closure and the perception of shape

James H. Elder

Department of Electrical Engineering

McGill University

Montréal

June, 1992

A thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Master of Engineering

Abstract

In the domain of visual perception, contour closure has been studied as a perceptual feature, extracted by preattentive visual processes. Motivated by basic results in topology and geometry, I advance a different hypothesis: perceptual closure should be seen not as a feature, but as a *means*; a dimensional bridge, continuous in nature, which allows the inference of planar and solid shape from one-dimensional contour. I support this hypothesis with a series of visual search experiments in which subjects discriminate outline figures by means of their two-dimensional shape. By modulating the degree of closure of the outlines, I show that two-dimensional shape processing is rapid for closed stimuli but slow for open stimuli. I further show that search speed can be characterized as a smooth, monotonic function of the degree of closure, supporting the notion of a perceptual closure continuum. This continuum is shown to be naturally parameterized by a dot-spacing metric, and experiments with different types of fragmentation lead to the hypothesis of a Minimax Principle of closure, which imposes a partial ordering over all fragmentations of a given contour. Using search speed as a metric for the perceptual closure of various figures leads to an equivalence relation between perceptual closure and the confidence of object inference. Under this metric, perceptual closure is found to be independent from various surface completion and texture processes and highly sensitive to contrast sign reversal, supporting the existence of a limited-complexity shape from contour process in early vision. Perceptual closure is shown to be nonlinear in several respects: a simple model of one such nonlinearity is proposed.

Sommaire

Dans le domaine de la perception visuelle, la fermeture du contour a été étudiée comme un trait perceptuel, extrait par des processus visuel préattentifs. Motivé par des résultats fondamentaux de topologie et de géométrie, j'avance une hypothèse différente: on doit voir la fermeture du contour non pas comme un trait, mais comme un *lien*: un lien dimensionel, ayant une nature continue, et permettant l'inférence d'une forme planaire et solide à partir du contour. Je soutiens cette hypothèse par une série d'expériences dans lesquelles des sujets distinguent des contours de figure sur la base de la forme bidimensionelle associée. Par la modulation du degré de la fermeture des contours, je montre que la distinction des formes bidimensionelles est rapide pour les stimuli fermés mais lente pour les stimuli ouverts. Je montre de plus que la vitesse de la distinction peut être caractérisée comme une fonction lisse et monotonique du degré de la fermeture, de ce fait supportant l'idée d'un continuum de la fermeture perceptuelle. Des expériences avec des fragmentations diverses mènent à l'hypothèse d'un principe de fermeture dit du "minimax". Ce principe impose un ordre partiel sur les fragmentations possibles d'un contour donné. L'utilisation de la vitesse de discrimination comme métrique pour la fermeture perceptuelle de formes diverses mène à une relation d'équivalence entre cette même fermeture perceptuelle et la confiance en l'inférence d'un objet. Sous cette métrique, la fermeture perceptuelle est trouvée être indépendante de la texture et des processus de la complétude modale et amodale, mais très sensible au changement de signe du contraste, supportant l'existence d'un processus de la vision primaire qui infère des propriétés de la forme bidimensionelle. Je montre que la fermeture perceptuelle est non-linéaire par certains égards: un modèle simple d'une propriété non-linéaire est proposé.

Acknowledgements

I thank all 29 of my hapless subjects for allowing me to probe their inner souls: they certainly didn't do it for the money, and will have to be content with fame and glory.

Many of my colleagues have helped considerably in reviewing this work or previous versions of it. They include Allan Dobbins, Greg Dudek, Thomas Jelonek, David Jones, Michael Kelly, Michael Langer, Frederic Leymarie, Anna Lin, Ron Rensink and Frances Wilkinson.

It is customary to thank one's supervisor for feedback and guidance. I would like to take these as assumed, and save my thanks for something less common. The computer vision lab of McRCIM is provocative and fun: a place where one can be passionate about ideas. In my opinion, this is due in large part to Steve Zucker. His ideals provide a powerful and exciting structure in which to work, to define oneself for or against. I am lucky to be here.

Deepest thanks are saved for my parents, Harvey and Marian, for their support, their love, their ethics and ideals.

Table of Contents

List of Figures

List of Tables

Chapter 1 Overview

Amongst the many contours in an image, there will be some which project from the boundaries of objects. Integrating information from these occluding contours allows the inference of two and three-dimensional shape properties. The trick is to pick the right contours to integrate, so that non-occlusion contours, or occlusion contours from distinct boundaries, are not integrated into representations of non-existent objects.

The hypothesis motivating this work is that this process of selective integration is based upon contour closure. This perceptual closure is presumed to have some correspondence to mathematical and intuitive notions of closure, but also to have properties specific to a perceptual context.

This hypothesis is based on the fact that the boundary of an unoccluded object with a simply connected surface projects as a simple, closed contour in a retinal image. A classical result in topology known as the Jordan curve theorem states that such a contour partitions the plane into two sets: an inside (figure) and an outside (ground). This partition is crucial because it allows the definition of two dimensional shape properties such as curvature sign, concavities, convexities, narrowings and bulges. Such properties have in turn been shown to impose strong constraints on the shape of the three-dimensional surfaces in the scene (Biederman [1988]; Koenderink [1984]; Lowe [1985]; Marr [1982]). Thus, computationally, notions of contour closure and shape are deeply entwined: contour closure makes the inference of shape from contour possible.

While in topology there is no concept of partial contour closure (a curve is either closed or it's not), in perception, shape can still be perceived even when the bounding contour is fragmented by occlusion, shadow or low reflectance contrast. What then is the role of closure in the perception of shape?

My hypothesis is that in perception the notion of a *closure continuum* exists and that it is this *perceptual closure* which mediates shape perception. This hypothesis connecting closure to shape has never been tested: contemporary research has instead divorced these two notions, characterizing closure as an "emergent feature", "completely abstractable from shape" (Treisman & Paterson [1984]).

I support this hypothesis with a series of visual search experiments in which subjects discriminate outline figures by means of their two-dimensional shape. The stimuli are designed so that the closure of the outlines can be varied without changing the contour features which distinguish the target from the distractors. By modulating the degree of closure of the target and distractors in tandem, I show that two-dimensional shape processing is rapid for closed stimuli but slow for open stimuli. I further show that search speed can be characterized as a smooth, monotonic function of the degree of closure, supporting the notion of a perceptual closure continuum.

Using search speed as a metric for the perceptual closure of various figures leads to a number of properties of perceptual closure and several principles governing the perception of shape from contour.

Experiments using partly closed figures of various fragmentation geometries lead to the hypothesis of a *Minimax Gap Principle*: given a length of contour and a region to be bound, maximum closure is obtained by minimizing maximum gap length. A simple mechanism based upon nonlinear operators is proposed as a possible basis for this principle.

Experiments with contrast-reversing contours lead to a *Contrast Sign Principle* of closure: perceptual closure is a function only of contour of a consistent contrast sign. This principle associates perceptual closure with other short-range contour grouping processes and distinguishes it from more complex, long-range surface inference processes, supporting the role for closure in early vision.

Experiments using first-order random-dot textures over the interior of fragmented outline figures reveal that large texture differences between figure and ground do *not* speed the formation of planar shape representations, suggesting a lack of cooperation between region and boundary processes at an early stage.

Further experimentation shows that modal or amodal completion of partly closed figures also does not speed shape discrimination, supporting the existence of a limited-complexity shape from contour process in early vision.

Figure symmetry is seen to affect discrimination in two ways. Reducing symmetry of highly closed figures is shown to slow discrimination *mildly*, reflecting the existence of a symmetry norm in the coding of shape. Reducing symmetry of figures with low closure is shown to slow discrimination *dramatically*, reflecting the exploitation of symmetry in perceptual grouping.

The majority of my results can be unified under a functional view of perceptual closure as a measure of the confidence with which contours can be interpreted as projections from a single object boundary. This leads to an equivalence relation between perceptual closure and *single object confidence.*

1.1 Contributions

- Countering the prevailing view of contour closure as a perceptual feature, I propose an alternate view which relates ideas in topology and geometry to perception, providing insight into the inference of multi-dimensional shape from contour.

- Through psychophysical experimentation, I support my hypothesis by demonstrating that contour closure is a strong determinant of rapid two-dimensional shape discrimination.

- I show that for simple figures, discrimination speed can be used as a metric for perceptual closure, and provide experimental support for the notion of a perceptual closure continuum mediating shape perception.

- I show that the perceptual closure continuum is naturally parameterized by a dot-spacing metric, and provide support for a Minimax Principle of closure which imposes a partial ordering over all fragmentation geometries of a given contour.

- I propose an equivalence relation between perceptual closure and the confidence of object inference, and provide experimental support for this relation.

- I demonstrate the sensitivity of perceptual closure to contrast sign reversal and the independence of perceptual closure from surface completion and texture mechanisms, and argue that these results provide evidence for the existence of a limited-complexity shape-from-contour process in early vision.

- As a methodological aside, I present evidence that the traditional psychophysical procedures used for visual search produce biased results, and demonstrate an alternate procedure that avoids this bias.

1.2 Things to See While in this Thesis

In the next chapter, I discuss the different sorts of contours one finds in images, and motivate the role of closure in the inference of shape from contour.

In chapter 3, I provide a brief review of previous research on the role of closure in visual perception, from the Gestalt years to the present.

The heart of the thesis is the Experiments chapter (chapter 4). In order to make it slightly less soporific, I have included some discussion of results with the presentation of each experiment. However, the pictures still tell most of the story, and the Table of Contents should help you to zero in on experiments of particular interest.

I strongly recommend the first section of chapter 5, as it wraps the results of many of my experiments into a functional view of closure as a measure of single object confidence. The second section distinguishes between the confidence with which an inference can be drawn, given a particular image, and the computation required to draw it. This distinction allows us to better understand the limits of the single object confidence hypothesis in predicting discrimination speed. The third section shows how limited simple topological and geometrical tools are for modeling perceptual closure. The metaphor in the fourth section is for fun.

The final section of chapter 5 discusses the various nonlinear aspects of perceptual closure revealed by my experiments, and proposes a simple mechanism to model one such aspect.

In the Conclusion (chapter 6), I briefly discuss the implications of my results for biological and machine perception, and suggest ideas for further experimentation.

In the appendix I validate my basic results using the standard visual search methodology, and present evidence that this methodology is subject to a systematic bias which my method avoids.

Chapter 2 Introduction

2.1 Contours

An important class of visual tasks (e.g. recognition, discrimination, manipulation, inference of function) requires the notion of a physical object and its shape. An *object* is some bounded structure in the world, typically with a significant set of common properties. While the properties which determine whether we call something one object or two are sometimes of a semantic nature (e.g. a cup and a saucer, since they function as a unit, may at times be considered one object), there are many properties, geometric in nature, which determine object under more general premises.

By the *shape* of an object I simply mean its spatial organization. Since most properties which we can sense are of the *surface* of an object, it is really the organization of this surface that we are interested in. The human visual system picks up surface properties using several basic sensory quantities: light intensity, colour, time, retinal position and multi-sensor disparity. In my experiments I restrict my focus to just two: intensity and position. That these alone are rich sources of information is evidenced by the value of black and white photography and drawing.

Difference encoding is one of the key principles of the retina and LGN. In the spatial domain, this is accomplished by means of center-surround cells, which compute a measure of local contrast. Cortical processing of visual information in primate brain begins in the striate cortex (V1), which receives a precisely organized projection from the LGN. While

LGN neurons are typically circularly-symmetric, most striate neurons exhibit oriented receptive fields, often quite narrowly tuned (Hubel & Wiesel [1968]). Thus, while cells in the retina and LGN are responsive to local change, striate cells respond to local, *oriented* change.

Oriented image change can arise from a variety of scene events. This thesis is concerned with the oriented structure of image contours, which project from space curves in a scene. These space curves and their projections are of three types: *occlusion curves* demarcate the visible portions of surfaces, *reflectance curves* outline surface regions which are distinct in some material properties, and *shadow curves* bound surface patches which are eclipsed by other surfaces in the scene. Note that while occlusion curves depend upon viewer geometry, and shadow curves depend upon light source geometry, reflectance curves are often independent of both. All three types of contour embody information about the surfaces in the scene.

Drawings which represent only occlusion contours are often excellent depictions of visual scenes (Fig. 2.1). Occlusion contours are intuitively important: they provide constraints on the extent of different objects in the scene, they give us a 'slice' of the surface shape of these objects, and through the inference of features such as T-junctions, they can inform us of the depth ordering of surfaces relative to the perceiver (Nitzberg [1991]).

Occlusion contours are also unique in the diversity of visual qualities by which they may be defined. While we may normally think of object boundaries projecting as luminance edges, occlusion contours are often defined by edges in disparity, velocity, texture and colour as well. Coincident change in several of these properties can thus serve as a means for distinguishing occlusion contours from reflectance and shadow contours.

(a) (b)

Figure 2.1: a: *Nude*, Henri Matisse (Clark [1956]). **b:** Contour drawing of figure from *Studies for the Libyan Sibyl*, Michelangelo (Nicolaïdes [1941]).

Reflectance contours are also important. Certainly the hair markings of giraffes, zebras and leopards are recognizable and useful for identification. Of course, you would not be able to read this sentence, nor appreciate Fig. 2.1 were you not able to represent and interpret shape from reflectance contours. The shape of shadow contour can also be informative, both about the shape of the surface being shadowed and the shape of the shadowing surface (Cavanagh & Leclerc [1989]).

This thesis is primarily concerned with occlusion contours, or depictions thereof, and how they can be used to infer multi-dimensional shape.

2.2 Bridging the Dimensional Gap

Typical scenes generate many contours in an image, and these contours are in some way encoded by the brain. But how much does this have to do with our perceptions and behaviours? Surely what we are most interested in is the character of surfaces and volumes in the world: should we not be trying to compute representations of these?

Indeed we should. Much of modern vision research, particularly computer vision research, has focused on methods for estimating relative depth and orientation of smooth surfaces in a scene, using measurements of motion, disparity, texture, shading, etc... (Marr [1982]). These methods have enjoyed some success in determining local surface geometries, at least within the assumptions (often severe) required by their methods, though they have contributed very little to our understanding of what is important about shape and how we might go about representing it (Witkin & Tenenbaum [1983]).

It does seem logical to analyse local properties of image patches which are projections of surfaces in the scene in order to infer surface shape. However, we need only glance again at the forms in Fig. 2.1 to be reminded of the powerful perceptions of three-dimensional shape which can be elicited by contour alone. Moreover, when occluding contour and shading information conflict, contour typically determines the percept (Biederman [1988]). Somehow our visual system *can* and *does* form excellent two- and three-dimensional percepts solely from one-dimensional contour information. How does it bridge this dimensional gap?

I will argue that a key step in this effort is the exploitation of *contour closure*. This hypothesis is based on the fact that the boundary of an unoccluded object with a simply connected surface projects as a simple, closed contour in a retinal image. A classical result in topology known as the Jordan curve theorem states that such a contour partitions the plane into two sets: an inside (figure) and an outside (ground). This partition is crucial because it allows the definition of two dimensional shape properties such as curvature sign, concavities, convexities, narrowings and bulges. Such properties have in turn been shown to impose strong constraints on the shape of the three-dimensional surfaces in the scene (Biederman [1988]; Koenderink [1984]; Lowe [1985]; Marr [1982]). For example, the sign of curvature of an occluding contour constrains the curvature of the projecting surface. If the contour curvature is negative (thus forming a concavity), then locally the surface must be hyperbolic (saddle-shaped). If the contour curvature is positive (forming a convexity), the surface is elliptic and convex. If the contour has zero curvature, the surface is parabolic (cylindrical) (Koenderink [1984]).

There is no question that we can perceive multi-dimensional shape from contour. Recent physiological results provide more food for thought: some neurons in area TE of primate cortex respond selectively to shape properties (concavities and convexities) which are strictly two-dimensional (Tanaka et al. [1991]). Mathematical results indicate that such computation depends upon the property of contour closure. Does this mean that the brain computes topological properties? If not, then what do we mean by contour closure in the domain of visual perception? This is the topic of my thesis.

Chapter 3 Background

3.1 Closure: A Principle of Organization

The importance of contour closure to perception was observed and discussed by Gestalt psychologists in the 1920s and 1930s:

> Ordinary lines, whether straight or curved, appear as lines and not as areas. They have shape, but they lack the difference between an inside and an outside... If a line forms a closed, or almost closed, figure, we see no longer merely a line on a homogeneous background, but a surface figure bounded by the line (Koffka [1935]).

This short excerpt captures the Gestalt view of closure. Therein we find the distinction between 1-dimensional and 2-dimensional ('surface figure') shape. We find a clear translation of the Jordan curve theorem in the distinction of inside and outside, but at the same time an appreciation of the importance of *partial* closure.

Fig. 3.1, taken from the same source, is one of the classical illustrations of the Gestalt theory of perceptual organization, illustrating in this case the predominance of closure over proximity in determining the perceptual organization of the figure.

These ideas (the dimensional nature of shape, the appreciation of the Jordan curve theorem in a perceptual context, the concept of *partial* closure, the view of competing and cooperating factors) form the basis of my work. It is, however, at least as important to see the bounds of the

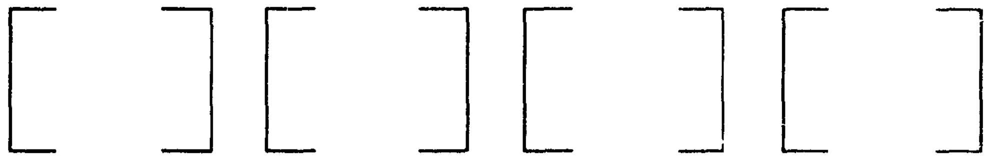


Figure 3.1: Gestalt illustration of the principle of closure

Gestaltists' conceptions as it is to see their extent. In particular, evidence of their loosely-stated claims about closure was limited entirely to introspective evaluations of a single figure (Fig. 3.1). What is the behavioural relevance of these ideas? Could these factors generalize beyond simple drawings, or even beyond this one simple figure? Was there any way to more precisely state these principles? To quote Koffka himself

> ...I should be the last to be satisfied with my hypothesis. Not only does it, as yet, lack experimental evidence, but it is not explicit enough, it contains no statement about the actual forces along the contour line... (Koffka [1935])

This thesis takes up the challenge of experimentally determining the role played by closure in the perception of shape from contour. First, however, we must jump ahead another 50 years to the context of contemporary perceptual research. For although the spirit of my work springs more from the Gestalt period than our own, the psychophysical techniques and, perhaps more importantly, the desires for more precise (read ***exploitable***) models are definitely modern.

To my knowledge, research on contour closure in perception lay dormant until quite recently. In the last fifteen years, however, several different labs have published relevant experimental results. While each of these labs has its own set of psychophysical tools and specific models, the similarities in their approaches and conceptions, particularly when considered in the context of their Gestalt precursors, are far more striking than their differences. Thus, while I will briefly outline their respective results with respect to closure, my primary goal here is to clarify the current conventional wisdom, as defined implicitly by these labs, and to bring that into relief against the original Gestalt ideas and my own work to be described.

The approaches taken by these labs are greatly influenced by two major contemporary concepts of vision research: those of *features* and *attention*.

3.2 Features and Textons

The major inspirations for modern feature theory were early physiological results revealing properties of the visual systems of animals such as frog (Lettvin et al. [1959]), cat (Hubel & Wiesel [1960b]; Hubel & Wiesel [1962]; Kuffler [1953]) and monkey (Hubel & Wiesel [1960a]; Hubel & Wiesel [1968]). These experimental results were exciting as demonstrations of the highly structured nature of visual processing, in which different neurons and fibres encode different attributes of a visual scene. The fashion in perceptual organization research has been, however, to extend and simplify these results to declare the existence of primitive feature detectors which determine our ability to rapidly discriminate small figures or textures.

This view holds that the first stage of human visual processing is a rapid, spatially parallel transformation of image information into a small number of independent, retinotopic feature maps. Each of these maps registers the presence or absence of simple image features at every location in the optic array. The primary goal of these researchers is to identify just what these features are.

This is of course a caricature of a whole field of research, and the papers I will discuss deviate somewhat from the canon. For example, Pomerantz et al. (Pomerantz, Sager & Stoever [1977]) entertain the notion that different features may not be completely independent, in that one may be used in the computation of a second, "emergent" feature. Julesz (Julesz [1991]) considers his texton theory independent of Treisman's feature theory (Treisman & Gelade [1980]), because he uses more crowded visual displays in his experiments and is motivated by the problem of segmenting an image into projections of different surfaces based upon differences in image texture. The elements making up Julesz' textures, therefore, are presumably depictions of surface markings. Treisman's visual search experiments, on the other hand, are motivated by the problem of finding some target object in a jumble of distractor objects. The stimuli used in these experiments could therefore sometimes depict occlusion contours of separate objects in a potential scene.

Despite these differences, the basic hypothesis (rapid, early transformation to maps of simple primitives) is identical across these theories. For all of these researchers, the prime research question pertaining to closure is: is closure a perceptual primitive?

3.3 Attention

There is ample motivation for the concept of attention in perception. We clearly have an introspective understanding of attention, since we can change our awareness of different objects in the visual field without changing fixation. Since so many things of interest in our world are localized in space and time, the ability to transfer computational machinery to different locations at different times is highly useful.

The modern view is that the human visual system can be partitioned into a "pre-attentive" and "attentive" system (Neisser [1969]). The preattentive system is responsible for the parallel transformation of image information into feature maps. The attentive system does everything else, and it does it in a serial fashion.

While the concept of attention pervades modern research in perceptual organization, the precise terms used vary. Pomerantz used the terms *parallel* and *serial* rather than preattentive and attentive (Pomerantz, Sager & Stoever [1977]). Julesz has used terms such as *spontaneous* (Julesz [1962]), *without scrutiny* (Julesz et al. [1973]) and *effortless* (Caelli & Julesz [1978]), all corresponding roughly to the term *preattentive.* Ullman refers to *early visual processes* and later *visual routines.*

The concept of attention is crucial as the link between feature theory and psychophysics. If two textures cannot be discriminated within a certain time limit, or if the time to find a target stimulus depends upon how many distractors are in the display, it is assumed that the stimuli or texture elements cannot differ in their primitive features. Otherwise, the preattentive system could have been used to rapidly solve the task. The feature map/attention story thus forms a basis for the interpretation of visual search and texture discrimination results. I will briefly discuss some of the results that pertain to the subject of contour closure.

3.4 Pomerantz

In their 1977 paper, Pomerantz et al. set out to test the effect of stimulus context on the discrimination of simple contour drawings. Their inves-

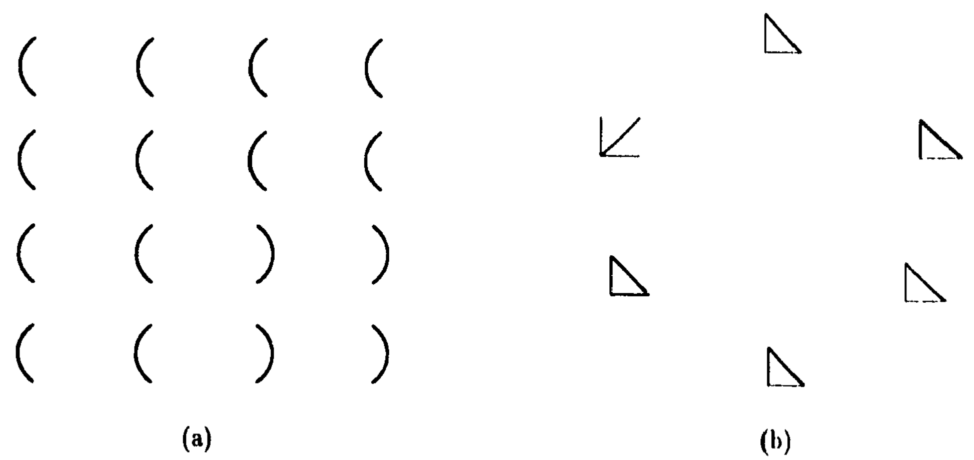


Figure 3.2: a: Square array for oddity task; b: Circular array for oddity task

tigation was based on what they called an 'oddity task'. Subjects were presented with a sequence of displays containing a number of stimuli arranged either in a regular square array (Fig. 3.2(a)), or regularly spaced around an imaginary circle (Fig. 3.2(b)). For experiments based on the square array, subjects were to identify the quadrant in which the stimuli were different. In the experiments using the circular geometry, subjects were to respond if and as soon as an odd stimulus was detected in the display.

Pomerantz et al. found that adding identical stimulus information to both the odd stimulus and the background stimuli could either speed, slow or have no significant effect on discrimination, depending upon the nature of the original stimuli and the additions. They took greatest interest in those cases where adding the information speeded detection.

For example, they found that adding 'mates' to the contours of Fig. 3.3(a) to form the stimuli of Fig. 3.3(b) substantially speeded discrimi-

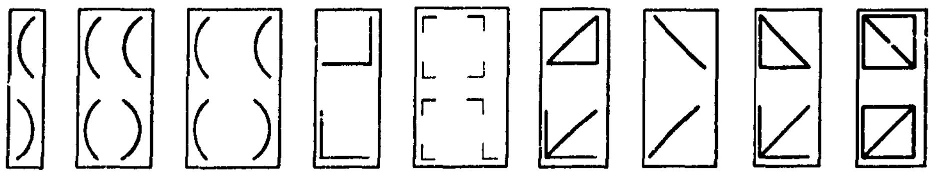


Figure 3.3: Selected stimuli relevant to contour closure

nation. This effect could be reduced by moving the two contour segments apart, as shown in Fig. 3.3(c).

In further experiments, they found that discrimination of the corner stimuli of Fig. 3.3(d) could be greatly speeded by the addition of several regularly arranged corners (Fig. 3.3(e)) or a single diagonal line (Fig. 3.3(f)). Similarly, the discrimination of two orthogonal line segments (Fig. 3.3(g)) could be enhanced by adding a corner (Fig. 3.3(h)). To balance these examples of what they called a "configural superiority effect", they showed many counterexamples where the addition of context significantly slowed discrimination (Fig. 3.3(i), for example).

In a final experiment, Pomerantz et al. used the circular display geometry to test the dependence of discrimination speed on the number of stimuli in the display, using the orthogonal line segments with no context, 'good' context or 'poor' context added (Fig. 3.3(g), (h) and (i) respectively). They found that while discrimination with no context or good context depended only mildly on display size (slope = 2.3 ms/item), with poor context this dependence was quite significant (slope = 76 ms/item).

While Pomerantz et al. did not invoke the concept of attention by name, they interpreted results by positing two extremes of processing. At one extreme, they proposed that subjects could be processing the displays in a *serial* fashion, one stimulus at a time. At the other extreme, subjects could be processing the stimuli in parallel.

They concluded that in some of their experiments, discrimination was performed serially, and in other cases in parallel. In particular, they claimed that the addition of good context could change a serial task into a parallel task.

These experiments cause problems for feature theories which assume that features are computed independently. The results of these experiments clearly show that the context of a local stimulus property greatly determines our ability to discriminate it, recalling the classical Gestalt conclusion:

> We could solve no problem of organization by solving it for each point separately, one after the other; the solution had to come for the whole... It has been said: The whole is more than the sum of its parts. It is more correct to say that the whole is something else than the sum of its parts, because summing is a meaningless procedure, whereas the whole-part relationship is meaningful (Koffka [1935]).

Pomerantz et al. attempted to resolve these problems by proposing the concept of "emergent features" which might be either local (vertices and intersections, for example), or global (symmetry and *closedness*). They argued that "wholes are perceived by these emergent features which are not the parts themselves but rather stem from the interaction of these parts".

Thus the concept of contour closure is reintroduced to perception as a potential "emergent feature" computed in parallel by the brain.

3.5 Julesz

In the early 1960s, Bela Julesz published results of discrimination experiments using random textures designed to minimize the familiarity of the stimuli (Julesz [1962]). Julesz argued that displays generated from a stochastic process would be "devoid of familiarity clues" and would thus "reveal some basic organization principles of information processing in the sensory nervous system".

In this early work, Julesz characterized textures in terms of their Nth-order statistics, defined as "...the probability that the vertices of an 'N-gon' (e.g. a hexagon, pentagon, etc...) thrown randomly on a texture fall on certain N colors." (Julesz [1991])

While one of Julesz' goals at this time was to study textures with identical N-th order statistics and to determine "the highest N that still yielded texture segmentation" (Julesz [1991]), the conclusions drawn in the original paper (Julesz [1962]) are actually of a local nature, in contrast to the global statistical approach taken in generating the figures: "...discrimination was found to be based primarily on clusters or lines formed by proximate points of uniform brightness."

Further experimentation revealed that this uniform brightness constraint could be relaxed to what may be best termed a 'uniform contrast sign' constraint, where sign is determined relative to the mean luminance of the texture.

Julesz described discrimination based upon these local properties as "connectivity detection", and went on to propose that texture discrimination is based on the existence and local properties (size, orientation, density) of locally connected regions of the textures.

In spite of the fact that the hardest conclusions from Julesz' 1962 paper were concerned with local properties of textures, what seemed to stick was the fact that *most* of the textures devised by Julesz that were identical in their 2nd-order statistics were difficult to discriminate. Thus, for example, mirror inverses are usually hard to discriminate, since any pattern is identical to its mirror inverse in its 2nd-order statistics, though it may differ in its 3rd-order statistics.[1]

Eleven years later, Julesz published another paper providing examples of textures which are nearly identical in their 2nd-order statistics and yet can be easily discriminated (Julesz et al. [1973]). In this paper, Julesz proposed a fairly specific two-stage model of texture discrimination. The first stage of this model consists of "early local feature extractors" that can detect only simple features such as dots and edges of given sizes and orientations. He modeled these feature extractors on "Kuffler units" and "Hubel and Wiesel units": that is, idealized models of receptive field structures found in the retina, LGN and cortex of cat and monkey.

The second stage of this model computes some global function based upon first- and perhaps second-order statistical differences in the outputs of these feature extractors.

In this paper Julesz did not characterize the hypothesized local feature extractors by psychophysical or theoretical means, he simply assumed they would have the idealized response properties of neurons found in early visual areas of cat and monkey. In two later papers (Caelli & Julesz [1978]; Caelli, Julesz & Gilbert [1978]), results of texture experiments were used as evidence for three such "perceptual analyzers", tuned to detect dots forming quasi-colinear, corner or *closed* structures, and serving as the "precursors of form perception". Caelli et al. referred

[1]This is due to the fact that a dipole and its reflection are related by a translation and a rotation, while a scalene triangle and its reflection are not.

to these as "Class B" detectors, to complement an hypothesized "Class A" set of detectors which distinguished differences in the second order (dipole) texture statistics.

Caelli et al. claimed that the nature of the Class B detectors indicated an underlying "visual geometry" which can bridge the processes of "effortless" texture perception and figure perception "with effort". In particular, they suggested that these Class B detectors could be the basic primitives initiating the discrimination of figure from ground.

Julesz soon revised this theory to postulate only two types of feature detectors: one for "elongated blobs" and one for "blob terminations" (Julesz [1981]). Thus he argued that preattentive discrimination of closed triangles from arrow patterns occurred on the basis of the line terminations, and not the property of closure itself.

Julesz claimed that two textures are preattentively discriminable only if they contain different numbers of a particular type of texton. Furthermore, he claimed that preattentive vision ignores the positional relationships between textons and the exact shape of the blobs, being sensitive only to their average width, length and orientation.

Although Caelli, Julesz et al. originally posited the existence of early visual closure detectors (Caelli & Julesz [1978]), the most recent statement from Julesz is that closure is not a texton (Julesz [1981]).

3.6 Treisman

3.6.1 Psychophysical Methods

Treisman et al. define a *dimension* as "a range of variation which is separately analyzed by some functionally independent perceptual subsystem (for example, shape, colour or orientation.)", and a *feature* as a value along that dimension (e.g. circular, red, or vertical) (Treisman & Gelade [1980]).

Treisman's Feature Integration Theory can be stated quite simply:

> The visual scene is analyzed at an early stage by specialized populations of receptors that respond selectively to such properties as orientation, colour, spatial frequency or movement, and map these properties in different areas of the brain... Features are registered early, automatically and in parallel across the visual field. Objects are identified separately at a later stage, requiring focussed attention (Treisman & Gelade [1980]).

Using these definitions, Treisman and her colleagues have tried to identify these primitive features using primarily two psychophysical techniques. In their visual search technique, subjects must find a target stimulus amongst several distractors. If the time required for a subject to find the target depends insignificantly or mildly on the total number of stimuli in the display, Treisman concludes that the target possesses a feature which none of the distractors possess. If, on the other hand, search speed depends strongly on the number of stimuli in the display, Treisman concludes that the target holds no such unique feature.

A second psychophysical technique is based upon Treisman's idea of attention as a spotlight which determines the accuracy with which features can be conjoined. In particular, if two or more items fall within the spotlight, their features could be interchanged, as could those of unattended items which fell outside the spotlight. She labelled these hypothesized phenomena *illusory conjunctions* (Treisman & Schmidt [1982]).

Treisman et al. demonstrated that when human subjects are shown visual search displays for very brief periods of time (120 ms, for example), and are asked to describe what they saw, they often make mistakes. They make even more mistakes if their attention is diverted.

The typical result of these illusory conjunction experiments is that subjects are more likely to make errors by incorrectly combining two features that are in the display than they are to make an error by combining one feature in the display with another feature not in the display. Thus features are not strictly tied to retinotopic locatic · they may "float" to conjoin with features at other locations.

3.6.2 Degrees of Closure

In (Treisman & Paterson [1984]), Treisman et al. set out to test the hypothesis that closure is an emergent feature. They performed five relevant experiments:

Experiment 1: Illusory Conjunctions of Lines and Angles

Examples of the displays used in this experiment are shown in Fig. 3.4. The experiment was performed in two blocks: in one block the target was an arrow junction, in the other a triangle. In both cases, the distractors were the right angles and tilted line segments shown in Fig. 3.4. Exposure duration was limited to between 45 ms and 270 ms.

Fig. 3.4 shows example displays in which the target is *not* present (Figures 3.4(a) and (b)), and displays in which the target *is* present (Figures 3.4(c) and (d)). A target was present for only 20% of the trials. Subjects were required to indicate whether they saw a target by pressing one of two buttons.

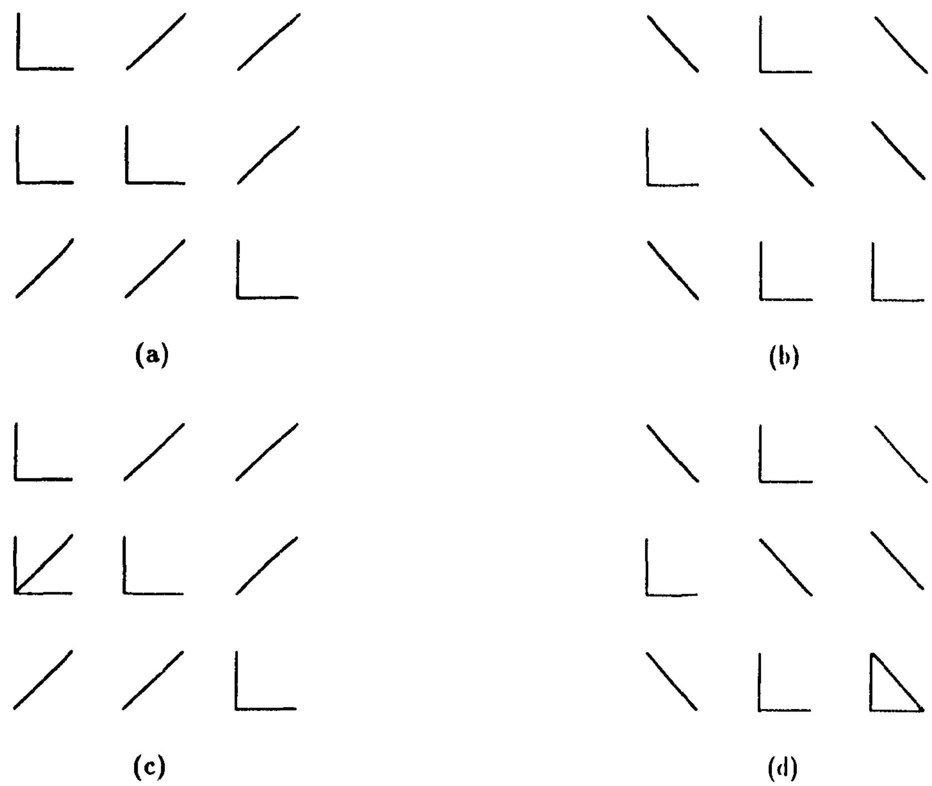


Figure 3.4: Stimuli used to examine the illusory conjunction of lines and angles.

The main result of this experiment is that subjects made more conjunction errors with an arrow target and like-oriented distractors than any of the other combinations, including that of the triangle target with like-oriented distractors.

Why was it harder to see an illusory triangle? Treisman argued that a real triangle possesses a feature that an illusory conjunction of angle and tilted line segment does not have: closure.

Experiment 2: Search for Arrow and Triangle Targets

A standard present/absent visual search experiment using displays similar to those shown in Fig. 3.4 revealed that search for the arrow target in like-oriented distractors was slower (7.7 ms/item for the target-present condition) than for the triangle target in like-oriented distractors (less than 2 ms/item for the target-present condition). Treisman uses this result to support her hypothesis that the triangle possesses a feature that the arrow lacks, allowing discrimination even against like-oriented distractor components.

Experiment 3: Texture segregation

This experiment used texture patterns formed from the stimulus items shown in Fig. 3.4. Treisman et al. found that the texture involving the closed triangles was more rapidly discriminated than that involving the arrow junctions. Again, they concluded that the triangles possess a distinctive feature that the other figures do not possess.

Experiment 4: The "Psychological Reality" of Triangle Lines

Illusory conjunction experiments were designed using the stimuli shown in Fig. 3.4, together with the character 'S', as distractors, and the symbol $ as target. The results showed that in fact all of the distractors: arrows, triangles and separated angles and tilted line segments produced illusory conjunctions with the S characters.

The authors concluded from these results that, while the angles and tilted line segments do not easily conjoin to form illusory triangles, the line components of triangles *can* conjoin with other display elements to produce new figures. They used these results to support their idea of

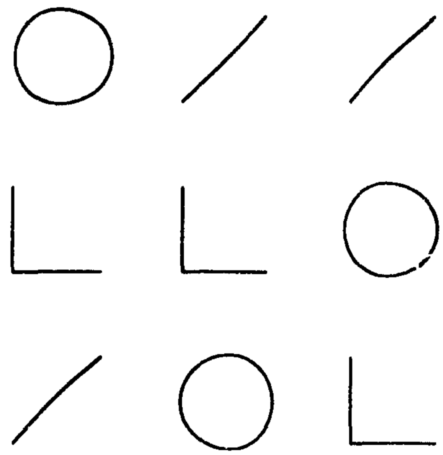


Figure 3.5: Display used to test the formation of illusory conjunctions from angles, lines and closed contours.

a separate, emergent closure feature, and to argue against a "holistic" version of perception, in which the elements of the triangle are "glued" together and thus prevented from conjoining with other elements in the display.

Experiment 5: Illusory Triangles from Angles, Lines and Closure

In this experiment, Treisman again used the illusory conjunction paradigm to see if adding closed contours to the displays containing angles and lines could increase the incidence of illusory conjunctions into closed triangle figures. An example of the displays used is shown in Fig. 3.5.

The results show that adding the closed circles to the displays *does* increase the incidence of closed triangle conjunction. The authors use this fact as further evidence for the existence of closure as an independent feature.

A relatively high number of closed triangles were also reported when

S distractors were used. This led the authors to speculate on the graded nature of closure:

> ...the perceptual definition of *closure* does not require *connected* lines but simply a space that is partially enclosed by a convex contour. There may be degrees of closure rather than all-or-nothing presence or absence.

3.6.3 Search Asymmetry

Given two different stimuli A and B, search speed for a target of type A amongst distractors of type B is often very different from that for a target of type B amongst distractors of type A (Treisman & Souther [1985]). Of most relevance to the topic of perceptual closure are the asymmetries found between open arcs of circles and closed circles (Treisman & Gormican [1988]; Treisman & Souther [1985]): while search for an arc of circle amongst circle distractors is parallel (does not depend significantly on display size), search for a circle in arcs is serial. Of even greater interest is the fact that the asymmetry in search speed increases as the arcs of circle are made more closed. Treisman saw this result as additional support for the notion of degrees of closure:

> ...line ends marking the gap are preattentively detected in parallel, whereas line connectedness is not. Instead, when the target is a closed circle, the relevant dimension appears to be a continuous one - degree of closure - that is shared to differing degrees by the distractor circles that have gaps (Treisman & Gormican [1988]).

Treisman has thus concluded that contour closure is a perceptual feature, extracted by a preattentive visual system, but that it differs from other features in its continuous quality.

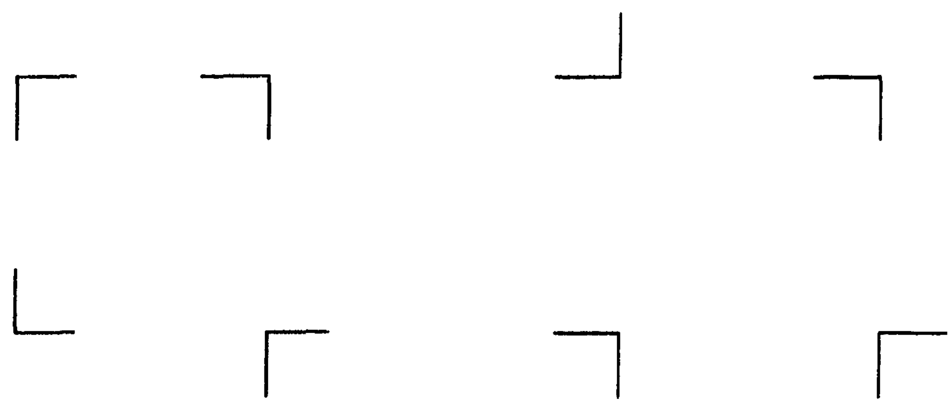
3.7 Donnelly

The effect of closure and good continuation on the rapid grouping of contour fragments into global object descriptions has recently been investigated (Donnelly, Humphreys & Riddoch [1991]). Donnely et. al. performed visual search experiments in which each bent contour fragment was considered a separate stimulus, and the fragments were arranged not randomly across the display, but regularly along a circular path centered on a fixation point.

In one experiment, the fragments were arranged so that the orientation change in each distractor fragment was toward the centre of fixation, and for the target was away from the centre of fixation (Fig. 3.6(a)). In this case, the time required to find the target was independent of the number of fragments in the display.

If, on the other hand, the distractors were arranged to change orientation away from the centre, and the target toward (Fig. 3.6(b)), search speed depended *strongly* on the number of fragments.

The authors claim that their findings suggest a parallel, preattentive mechanism for integrating features into combined object descriptions. They argue that in the first experiment, the distractor fragments are positioned and oriented to form a composite structure with the emergent properties of closure and good continuation. The target in this case is inconsistent (does not group with) this structure, and is thus selected as salient. In the second experiment, no such global structure forms, and



(a) (b)

Figure 3.6: The importance of global structure in determining search performance

the subjects are forced to consider the fragments individually to find the target.

The results of a final experiment in which the subjects are required to find an inconsistent fragment which may be in one of two circular configurations lead the authors to conclude that objects can be "selected" only one at a time, though some preliminary object computation may be done in parallel. With their stimuli, they estimate that this object selection requires about 200 ms.

Pointing to the smooth degradation of search speed with fragment jitter, the authors conclude that 'objectness' is a graded characteristic:

> Perceptual structures vary in the degree to which they can be selected as single objects, according to the degree of closure, good continuation and so forth... The graded nature of object coding allows flexibility of selection under non-optimal viewing conditions, when partial or incomplete visual infor-

mation is present.

3.8 Ullman

While debate has raged over whether to award closure the coveted feature status, very little thought has been given to how closure might be computed. One exception can be found in Shimon Ullman's Visual Routines paper (Ullman [1984]).

Ullman argues that certain operations, such as determining whether a point is inside or outside a closed curve, or determining whether a curve is in fact closed, are "essentially sequential": that is, they *require* a sequential procedure for their solution. For the inside/outside problem he suggests two procedures. The first involves drawing a ray from the point in question to infinity, and counting the number of intersections of the ray with the contour. If the number of intersections is odd, the point is inside the closed contour. If it is even, the point is outside.

Observing that this algorithm is limited to images consisting of a single closed curve, Ullman suggests a "colouring" or "bounded activation" scheme as an alternative. In this scheme, a colour is spread from the point in question to neighbouring points in the image. Colour spreading is continued subject to the condition that spreading cannot occur across a contour in the image. After some time (a practical representation of eternity), a far-away point in the image (a practical representation of infinity) is checked. If it is coloured, then the original point is *outside* the closed contour. Otherwise, it is inside.

To determine closure, Ullman suggests a contour marking and tracing routine. If in tracing a contour a termination point is encountered, the curve is open. If the starting point marker is encountered, the curve is closed.

This is a pretty naïve scheme, and it should become clear from my experiments that it has little to say about perceptual closure. I will return to the subject of visual routines in the concluding chapter of this thesis.

3.9 A New Perspective

It may be tempting to conclude that the most significant difference between the Gestalt movement and the modern approaches just described lies in methodology. The Gestaltists relied primarily on qualitative subjective impression. Treisman's work, on the other hand, emphasizes quantitative behavioural experimentation.

There is actually a much deeper difference. In seeking to understand perception, Gestalt psychologists sought *principles* which govern how we put image information together to form global percepts. Proximity, similarity, good continuation, closure, symmetry, *prägnanz*: these terms suffer from a lack of precision, but we cannot deny that, through inspection of the beautiful illustrations from this period, we understand at some level what these labels mean to perceptual organization. These are the forces that guide the construction of percepts. The term *forces*, while sadly associated with brain field theories, is apt because it captures the notion that these different principles are not binary rules, but smooth influences that cooperate and compete in determining consistent interpretations.

The modern approach, well represented in Treisman's Feature Integration theory and Julesz' Texton theory, is very different. They are *atomistic* as opposed to *mechanistic*. They seek to discover a "small set of simple, independent features" (Treisman & Gormican [1988]) or "perceptual quarks" (Julesz [1991]) upon which higher percepts are based. While the Gestalt psychologists accounted for the richness and subtlety of perception with the concept of continuous effects, these modern feature theories seek singular "atoms".

These differences are no better exemplified than in the study of contour closure. Julesz sought to answer the question "Is closure a texton?". Treisman wondered, "Is closure a feature?" Answering these questions through psychophysical experimentation is complicated by the fact that presence or absence of closure is never the only difference between the target and distractor stimuli. Thus, while some researchers have concluded from these experiments that closure is a feature (Caelli, Julesz & Gilbert [1978]; Pomerantz, Sager & Stoever [1977]; Treisman & Gormican [1988]; Treisman & Paterson [1984]), others claim that line terminations are the salient features, and it is only their absence which registers closure (Julesz [1980]; Julesz & Bergen [1987]).

These questions have little to do with the line of enquiry followed by Gestalt researchers, who sought to understand how closure could affect the organization of perceptual information and the definition of figure.

I take a view much closer to the Gestalt view. I will not try to even define the term "feature", and I certainly will not try to claim that closure is or is not such a beast. Instead, I will explore the role that contour closure plays in determining our perceptions of shape. In this view, closure is not seen as a *feature*, but as a *means*, involved in the construction of representations of shape.

Earlier, I motivated the hypothesis that closure plays a role in our perceptions of objects by the fact that the boundaries of these objects project as closed contours, and by topological and mathematical arguments which relate this property of contour closure to properties of shape.

But while in topology there is no concept of partial contour closure (a curve is either closed or it's not), in perception, shape can still be perceived even when the bounding contour is fragmented by occlusion. What then is the role of closure in the perception of shape?

My hypothesis is that a *closure continuum* mediates the perception of shape from contour. This hypothesis has never been tested: while Treisman has suggested the idea of a continuously coded closure feature, she has in her own work divorced the issues of closure and shape, characterizing closure as "completely abstractable from shape" (Treisman & Paterson [1984]).

I also use visual search techniques to explore properties of perceptual closure. In my experiments, however, discrimination is based not on the presence or absence of closure, but on the two-dimensional shape of the target and distractor. By modulating the degree of closure of both the target and distractor in tandem, I am able to characterize the influence of closure on shape perception, and to isolate certain properties of this perceptual closure.

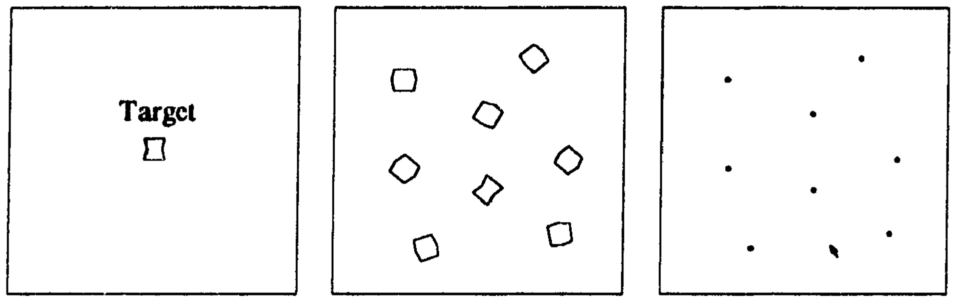
Chapter 4 Experiments

4.1 Methods

Visual search displays were created on a 60 Hz, noninterlaced colour Amtron monitor, driven by a Symbolics 3640 computer. Subjects sat in a dimly lit room, approximately 1 m from the screen. A $7^\circ \times 7^\circ$ square display window of luminance $11\,cd/m^2$ was positioned in the centre of the screen against a background luminance of $0\,cd/m^2$. Stimuli were normally drawn in the display window at $72\,cd/m^2$. Figure luminances used for experiments examining the effects of contrast reversal are specified in table 4.1 of section 4.14.

All stimuli were approximately $0.5^\circ \times 0.5^\circ$ in size, of random orientation and randomly jittered position (minimum inter-stimulus spacing $= 0.8^\circ$). In most experiments[1], displays contained either 7, 15 or 23 distractor stimuli and one target (display sizes of 8, 16 or 24 stimuli). The procedure is illustrated in Fig. 4.1. First, an example of the target for which the subject will be searching is shown (Fig. 4.1(a)). The subject then presses a mouse button to trigger a sequence of 30 visual search trials (10 for each display size, randomly interleaved). In each trial, a display is presented which always contains exactly one target (Fig. 4.1(b)). When the target is detected, the subject clicks a mouse button and the response time for detection is recorded. At the same time, the visual search display is replaced by a validation display in which the

[1] In some of the later experiments, only one display size is used. The changes in method are noted in the appropriate sections.



(a) (b) (c)

Figure 4.1: Visual search sequence

stimulus positions are represented by small reference dots (Fig. 4.1(c)). The subject must correctly identify the target location (by clicking on the appropriate dot) for the trial to be considered valid. If an error is made, the trial is rejected, and another trial with the same display size and stimulus type is randomly inserted in the sequence to replace it.

This procedure differs from traditional methods, in which half of the displays shown to each subject contain a target and half do not (Treisman & Gelade [1980]). In this procedure, subjects press one of two buttons, depending upon whether they perceive the target as present or absent. One advantage of my procedure is the relatively low error rates, which averaged 1.7% and were less than 5% for all experiments. More importantly, I believe that the traditional method is subject to a systematic bias which my method avoids.

I have reproduced the basic results of this work using the more traditional procedure. These results, together with an analysis of the problems with traditional techniques, can be found in the appendix.

Before each session, subjects complete a practice procedure identical to the recorded session, but including only 2 trials for each display condition. In addition, for every block in the recorded session, the first three trials (one for each display size) are used as practice, and the response times are not used in computing mean results, which are thus averaged over 9 trials for each stimulus/display size condition.

4.2 Subjects

A total of 29 subjects (21 male, 8 female) were used for the 40 experiments described here. Between 10 and 14 subjects were used for each experiment: the exact number is stated with the results. Subjects ranged from complete naïveté to full awareness regarding the goals of the study. All subjects had normal or corrected vision. Results are averaged over all participating subjects, with error bars indicating standard error of the mean.

4.3 Shape Discrimination for Open and Closed Figures

The basic stimuli (Fig. 4.2(a)) are composed of two unconnected but nearby contours. The contour segments are the same for both the target and distractor, which thus differ only in how the segments are placed relative to each other. In the target stimulus, they are arranged to bend inward, forming a 'spindle' shape, while in the distractor stimulus they bend outward, forming a 'barrel' shape. The stimuli have thus been chosen so that the discrimination must be based on measurements which are two-dimensional and extrinsic to the contour segments. That is, information from the two contour segments forming each stimulus must

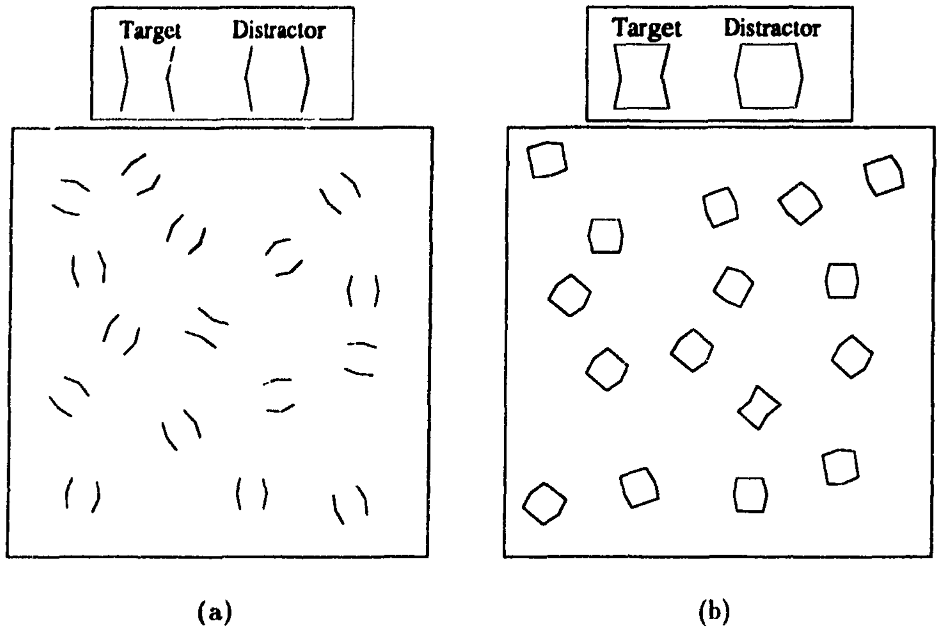


Figure 4.2: Stimuli and example displays for open and closed outline shapes.

be grouped into a composite representation which can then be used to discriminate the target from the distractors.

With the addition of two identical line segments to each of these open figures, two closed figures are formed (Fig. 4.2(b)). The length and relative spatial position and orientation of these segments is the same for each stimulus, thus by themselves they provide no direct means for discrimination or identification. Note that closing these figures endows them with new two-dimensional shape properties. For example, the closed spindle possesses two concavities and a narrowing which the closed barrel does not possess. These properties play a large role in computational theories of planar shape perception (Blum [1973]; Hoffman & Richards [1985]; Kimia, Tannenbaum & Zucker [1990]; Leyton [1989]) and so it is natural

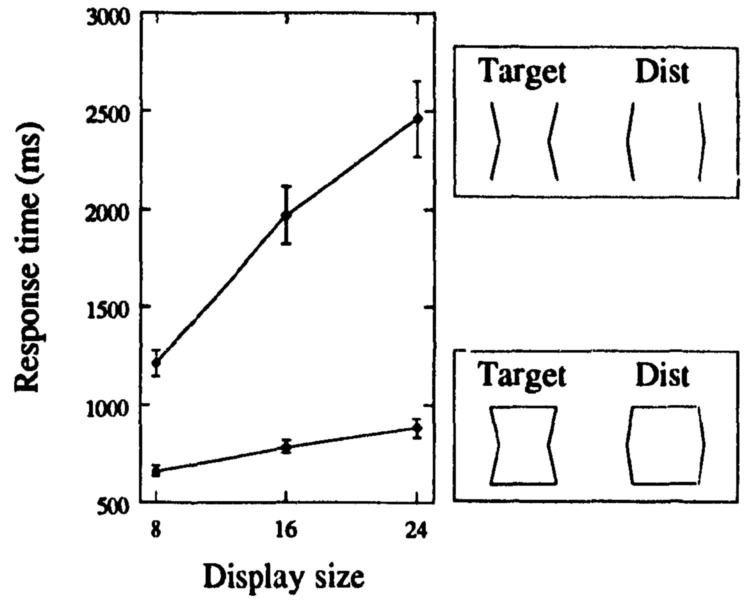


Figure 4.3: Search results for closed and open figures, averaged over 14 subjects.

to wonder whether closure, by making these properties well-defined, will make these shapes easier to discriminate.

The results show that closing the curves dramatically enhances discriminability (Fig. 4.3).[2] While search speed for the open shapes depends strongly on the number of stimuli in the display (slope = 83 ms/item, intercept = 555 ms), search for the closed shapes depends only weakly on the number of stimuli (slope = 14 ms/item, intercept = 546 ms), and is within the range of what is normally considered preattentive perception (Enns & Rensink [1991]; Julesz [1986]; Treisman & Gormican [1988]).

In order to draw solid conclusions, however, we must examine the other changes introduced by closing the stimuli. For example, the closed

[2]Slopes and intercepts were subjected to pair-wise one-tailed t-tests with a chosen significance level of 0.05. Closing the stimuli significantly reduced search slope ($p < 0.005$), but had no significant effect on the intercept estimate ($p > 0.1$)

contours possess four corners (orientation discontinuities), which could provide local information sufficient for discrimination. Also, the closed figures are connected. Connectedness has been proposed as an important rule of perceptual organization (Rock & Palmer [1990]): perhaps this property of connectedness is responsible for our results.

4.4 Corners, Connectedness or Closure?

Three experiments were performed which evaluate the importance of closure relative to other stimulus factors.

In the first of these (Fig. 4.4), the end quarters of the closing segments were removed to form stimuli possessing neither local corner information nor the property of connectivity. Although the stimulus contours are no longer topologically closed, intuitively they remain highly closed. The results confirm that while the absence of corners and connectedness does not null the effect, search performance is mildly degraded, consistent with a small decline in perceptual closure[3] (slope = 25 ms/item, intercept = 483 ms).

The next experiment further investigates the role of local information in this visual search task. Two pairs of stimuli were constructed, which both possess local corner information, but differ in their degree of apparent closure (Fig. 4.5). The length of each horizontal segment forming a corner is one quarter of the total gap size.

The results show that, while the local information is the same in both cases, there is an immense difference in subjects' ability to discriminate

[3]Removing the end quarters of the closing line segments resulted in a significantly greater search slope ($p < 0.025$), but had no significant effect on the intercept estimate ($p > 0.1$).

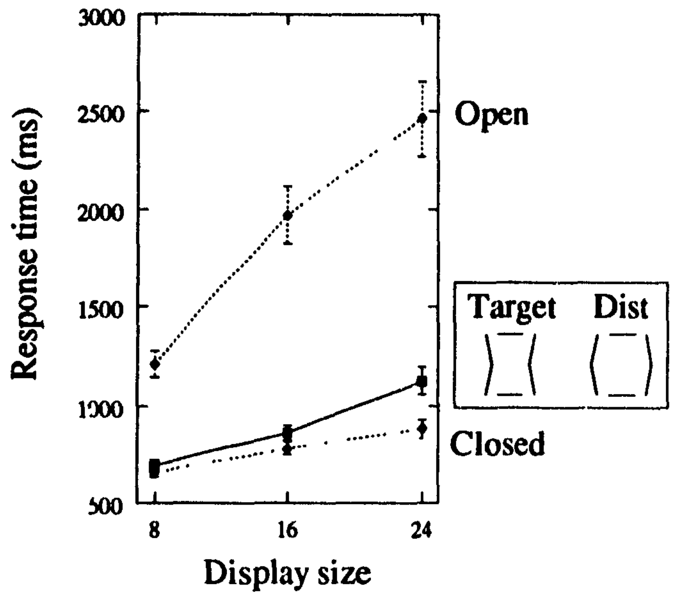


Figure 4.4: Search results for unconnected figures (14 subjects). The results for closed and open figures are shown dotted for reference.

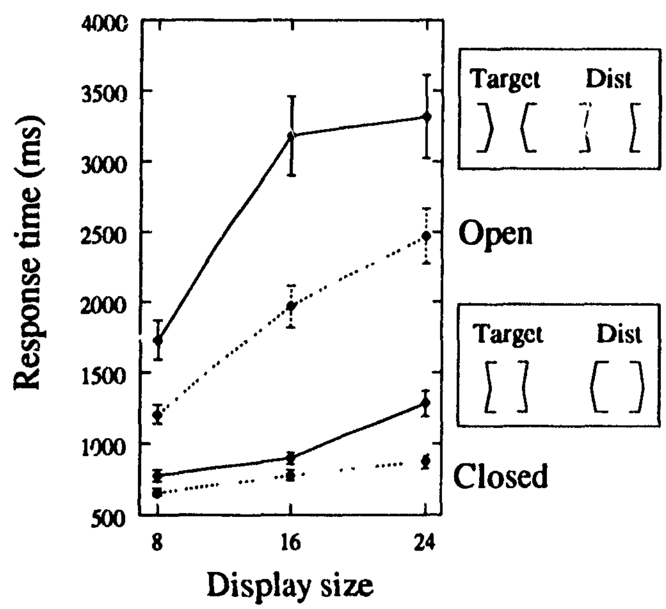


Figure 4.5: Search results for figures with inward or outward corners (14 subjects). Results for open and closed figures are shown dotted for reference.

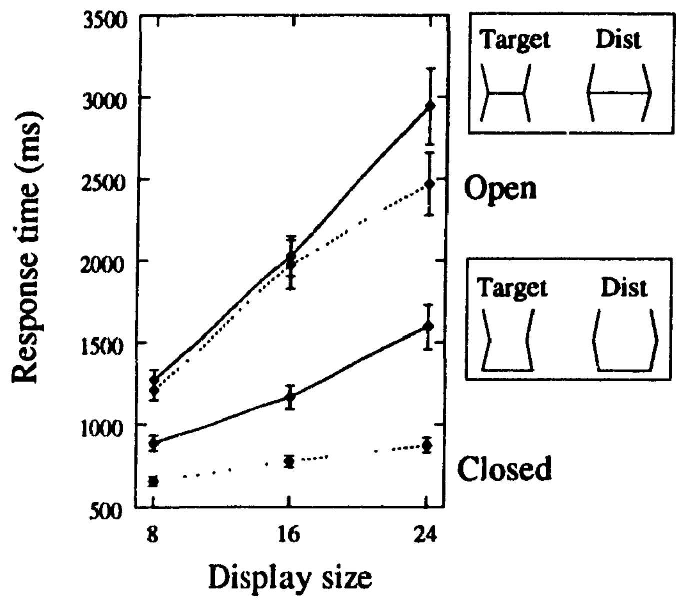


Figure 4.6: Search results for connected stimuli (14 subjects). Results for open and closed figures are shown dotted for reference.

the shapes.[4] When the corners were oriented inward to partly close the shapes discrimination was relatively rapid (slope = 27 ms/item, intercept = 535 ms). When the corners were oriented outward, discrimination was very slow (slope = 112 ms/item, intercept = 894 ms). Clearly, the global closure information is far more important than these local cues in forming the representations required to discriminate the stimuli.

A third experiment further investigates the difference between connectedness and closure (Fig. 4.6). Both experiments involve connected figures, but while the contours of one stimulus pair can be interpreted as partial object boundary projections (intermediate closure), those of the other pair are inconsistent with such an interpretation (poor closure).

Search speed for the stimuli closed at one end was intermediate be-

[4]Turning the corners outward resulted in a significantly greater search slope ($p < 0.005$) but had no significant effect on the intercept estimate ($p > 0.05$).

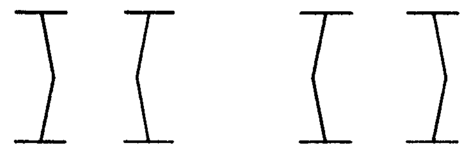


Figure 4.7: Inward and outward corner superposition

tween open and closed (slope = 40 ms/item, intercept = 556 ms), while search speed for stimuli connected at the middle was slow (slope = 100 ms/item, intercept = 466 ms).[5] Connectedness was not the dominant influence here.

4.5 Superposition

The second experiment of the previous section demonstrated that the orientation of corners in contour fragments has a dramatic effect on the closure of the composite stimulus. Inward corners result in a high degree of closure, while outward corners result in a low degree of closure. It is natural to ask the question: "What if the stimuli have both?" Figure 4.7 shows figures which do.

The results (Fig. 4.8) show that search for these stimuli is in fact significantly slower than that for both inward and outward corner stimuli (slope = 177 ms/item, intercept = 244 ms/item).[6]

[5]Search slope for the stimuli connected at one end was significantly greater than that for closed stimuli ($p < 0.005$) and significantly less than that for open stimuli ($p < 0.005$). Search slope for the stimuli connected at the middle was significantly greater than that for the stimuli connected at one end ($p < 0.005$), but not significantly different from that for open stimuli ($p > 0.1$). Intercepts for the end-connected and middle-connected stimuli do not differ significantly from each other or from those for the open and closed stimuli ($p > 0.1$)

[6]Search slope for the stimuli with both inward and outward corners was significantly greater than that for the stimuli with outward corners only ($p < 0.025$) and for the stimuli with inward corners only ($p < 0.005$) Intercept for the stimuli with

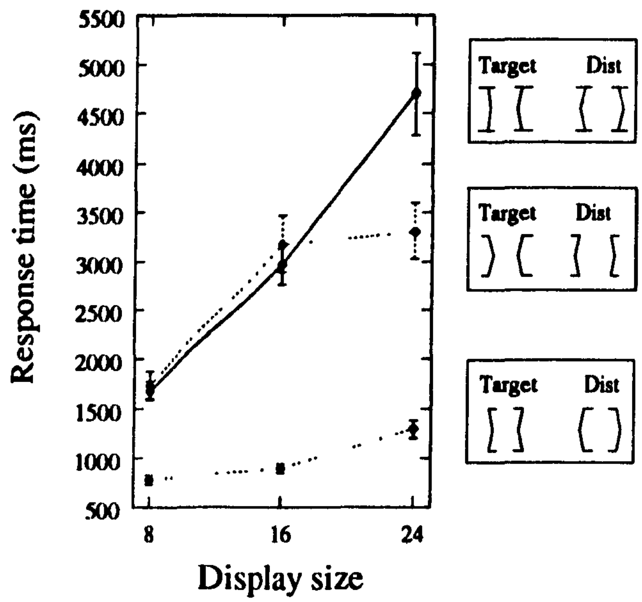


Figure 4.8: Results for junction stimuli (10 subjects)

While adding inward corners to the original open stimuli of Fig. 4.2(a) increased perceptual closure dramatically, adding inward corners to the outward-corner stimuli seems to reduce perceptual closure. This finding is consistent with a view of perceptual closure as a measure of the confidence with which the contour fragments can be inferred to project from the boundary of a single object. The contour fragments of Fig. 4.7 are not consistent with such an interpretation.

While the evidence for integrating the contour fragments composing the outward-corner stimuli is also poor, they could still be interpreted as fragments projecting from *two* distinct objects. In the case of the spindle fragments, these objects would be seen as convex, while for the barrel fragments, they would be seen as concave (Fig. 4.9). This would provide

both inward and outward corners was significantly less than that for the stimuli with outward corners only ($p < 0.05$), but did not differ significantly from that for the stimuli with inward corners only ($p > 0.1$).

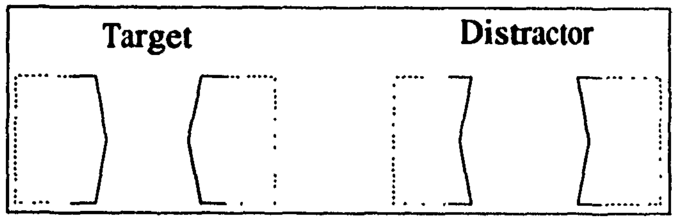


Figure 4.9: Interpretation of stimuli with outward-oriented corners

some two-dimensional basis for discrimination.

4.6 Competing Organizations

The visual search displays used in the above experiments consist of multiple stimuli in relatively close proximity to each other (Fig. 4.2). Perhaps the most basic of the original Gestalt laws of grouping is that of proximity: nearby image elements tend to group together. In our displays, if grouping was based upon image proximity alone, there might be cases where contour fragments of different stimuli would group together, forming composite representations that resemble neither the target nor the distractor. This confusion could plausibly lead to slower search, and in particular larger search slopes, for the incidence of such confusion can be expected to rise with the number and density of stimuli in the image.

It is therefore possible that the role of closure in these experiments is as a disambiguating image property, a factor which complements the influence of proximity in determining the perceptual organization of the image.

I have tested this idea in two ways. First, I repeated the original experiments with closed and open stimuli, but with a 50% increase in inter-stimulus spacing. The results (Fig. 4.10) show that substantially

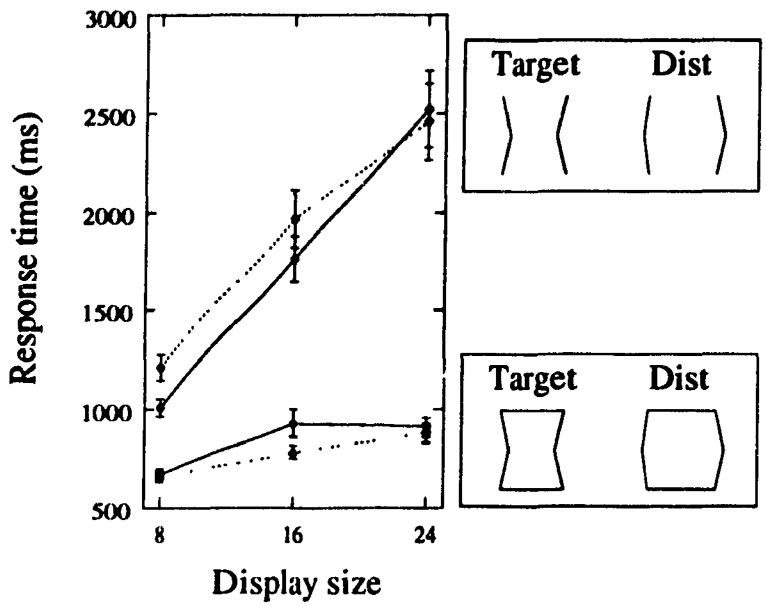


Figure 4.10: Search results for sparse stimuli (14 subjects). Results for closed and open stimuli at the standard density are shown dotted for reference.

decreasing the proximity of neighboring stimuli has no significant effect on search speed. (For the closed stimuli, slope = 16 ms/item, intercept = 548 ms/item. For the open stimuli, slope = 95 ms/item, intercept = 240 ms/item).[7]

While interpretation of this result is complicated by the simultaneous increase in total display size (from $7° \times 7°$ to $10.5° \times 10.5°$), it does suggest that closure has a larger role to play in perceptual organization than as a simple counteracting factor to proximity.

In addition to the traditional laws of grouping such as proximity and closure, Palmer has suggested a law of *common region* or *enclosure*: image features that belong to the same region tend to be grouped together

[7]Although there was no significant difference in search slopes ($p > 0.1$), increasing the inter-stimulus spacing did significantly decrease intercept for the open stimuli ($p < 0.05$).

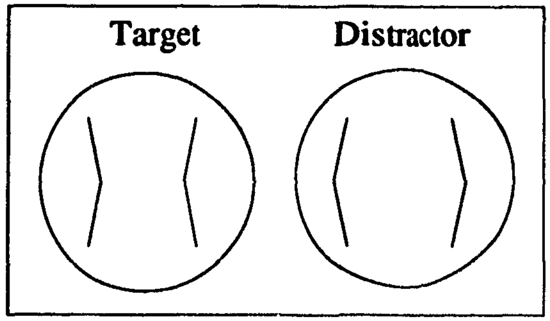


Figure 4.11: Open stimuli were enclosed by circles approximately 0.9° in diameter.

(Rock & Palmer [1990]). To test the importance of this property, I repeated the experiment for open stimuli but this time enclosed each stimulus within a circle (Fig. 4.11).

The results (Fig. 4.12) show that enclosing the open stimuli in this way has no significant effect on the parameters of a best-fitting linear model (slope = 101 ms/item, intercept = 532 ms/item). Although slopes and intercepts of a linear fit to the unenclosed and enclosed open stimuli do not differ significantly ($p > 0.1$), fitting a second-order model to the response curve for this enclosure experiment reveals a significant positive curvature ($p < 0.02$), and the best-fitting second-order models for the unenclosed and enclosed experiments are significantly different in all three parameters ($p < 0.05$).

Thus, although it appears that enclosing the stimuli certainly does not make the task easier, the addition of the circles does *change* the nature of the task: essentially making the task more difficult at larger display sizes. This may be due to the increased complexity of the display with the circles added.

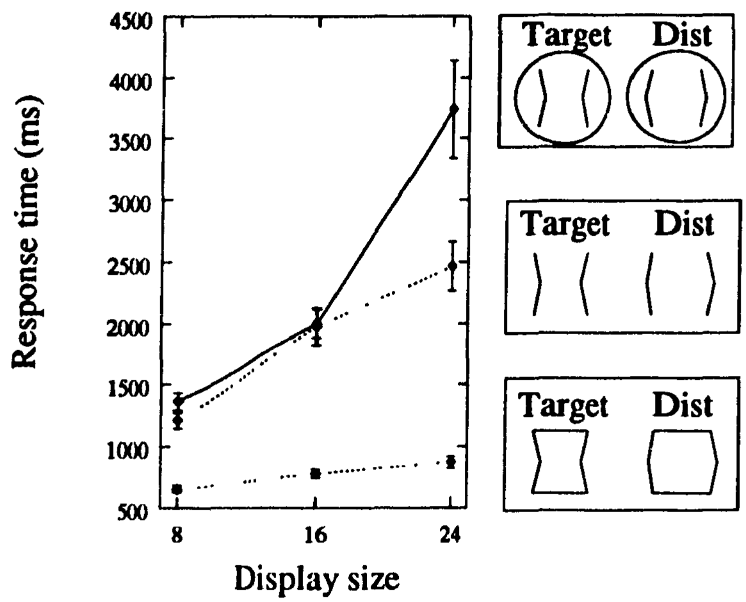


Figure 4.12: Search results for enclosed stimuli (14 subjects). Results for original open and closed stimuli are shown dotted for reference.

4.7 Is Closure a Floating Perceptual Property?

Treisman has suggested (Treisman & Paterson [1984]) that 'emergent properties' such as closure can 'float' and conjoin with other parts of an image. She shows (see section 3.6.2) that when asked to report on the presence or absence of a certain stimulus in a briefly-presented display, subjects sometimes incorrectly report it pres^nt. This tendency increases if the display contains elements which can be conjoined to form an instance of the target.

In these experiments, the task is made difficult by a very brief exposure duration (between 45 and 270 ms), a stimulus mask, and a distracting task which prevents subjects from attending to the stimuli. Under these stressed conditions it is not surprising that when forced to make a choice, subjects sometimes make mistakes by incorrectly conjoining features of different stimuli.

My experimental technique is very different. Subjects are not asked to report what they saw, but know in advance that a target is present and are asked to find it. If closure is a floating perceptual property, it is reasonable to expect that the closed circles of Fig. 4.11 would conjoin with the contour fragments enclosed, thus speeding search for this task. The fact that this does not happen suggests that we cannot think of closure as a floating perceptual property in this kind of task: it is a property *of* contour, and affects the way that contour is perceived.

4.8 The Perceptual Closure Continuum

The experiments of section 4.4 suggest that perceptual closure is not well-modeled as a topological property: when the bounding contour is fragmented, the ability of the human visual system to perceive two-dimensional shape is degraded but not destroyed.

To further clarify this, I created a new set of stimuli by incrementally adding contour to the original open barrel and spindle (Fig. 4.13). I conducted visual search experiments using these stimuli with a fixed display size of 16, characterizing search speed by the mean response time for each stimulus pair. The results are shown in Fig. 4.14, with response time plotted as a function of the number of pixels added to the open stimuli. Pixels added to form corners away from the figure are plotted in the negative abscissa range.

Intuitively, the stimuli in the high positive abscissa range have a very high degree of closure, while those in the high negative range have a very low degree of closure. The results of this experiment show that this intuition does correspond to the way closure is used to form rapid

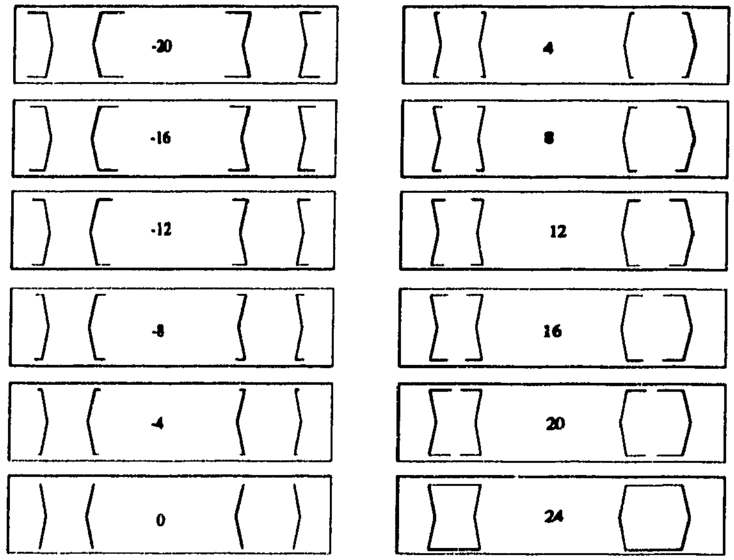


Figure 4.13: Sampling closure with 2-fragment stimuli.

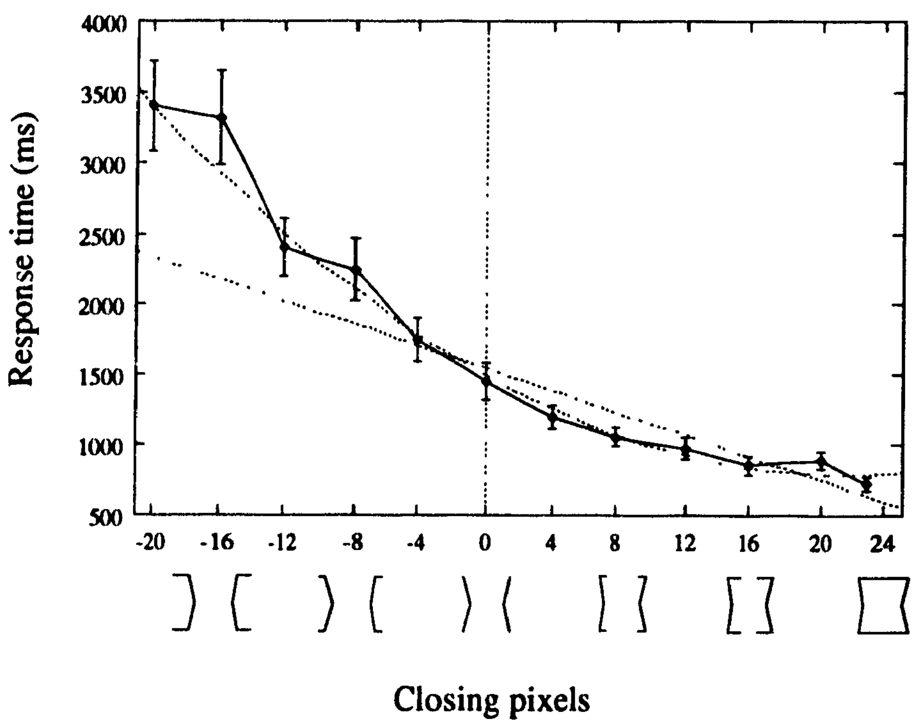


Figure 4.14: Linear and Quadratic fits to search results for 2-fragment stimuli (10 subjects, display size = 16).

representations of shape. For stimuli that we think of as highly closed, response time is rapid, and for stimuli that we think of as poorly closed, response time is slow. The concept of a continuum of perceptual closure is highly useful here.

A linear model of response time as a function of the length of the corner extensions is inadequate, however quadratic and exponential models both provide good fits.[8] Both of these models have positive curvature, indicating that contour fragment extension has a bigger impact on the absolute time required to perform the task when the degree of closure is small.

4.9 A Minimax Gap Principle of Closure

The previous experiment shows that as we add pixels to extend contour fragments so as to increase the closure of the stimuli, search steadily quickens. If we add pixels to extend contour fragments in a direction away from the figure, so as to decrease closure, search slows.

It is natural to wonder how much of this behaviour depends upon the way in which we extend the contour fragments of the stimuli. If we are adding pixels along the figure boundary, does it matter where we add them?

Two specific aspects of this question are:

1. Is it helpful if the pixels are added in a contiguous fashion, so as to

[8]The χ^2 for each fit was computed using the standard deviations for each stimulus condition. Each fit involved 1080 data points (10 subjects × 9 trials × 12 stimulus conditions). The linear model has 2 degrees of freedom ($y = a + bx$), the quadratic ($y = a + bx + cx^2$) and exponential ($y = a + be^{-cx}$) models have 3. The χ^2 for the linear model is 1120. Since the probability Q that, given a linear model, the χ^2 would exceed 1120 is 0, this is a poor model. The χ^2 for the quadratic and exponential models are 1074 and 1075 respectively, both yielding $Q = 1$

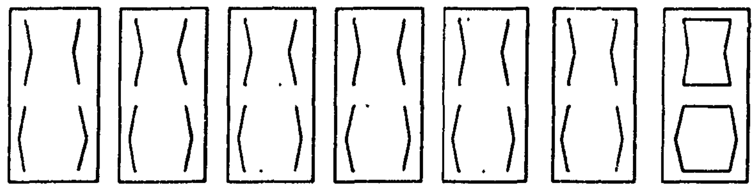


Figure 4.15: Boundary dot stimuli

establish or extend oriented structure along the figure boundary?

2. Is it helpful to add information in the vicinity of points of high curvature along the figure boundary?

The first question is partly motivated by the grouping law of connectedness proposed by Palmer (Rock & Palmer [1990]). The second question is motivated by models which stress the importance of information at points of high curvature (Attneave [1954]).

The stimuli shown in Fig. 4.15 were designed to shed some light on these issues. The stimuli were constructed by adding isolated, uniformly-spaced dots to the open stimuli along the boundary of the shape.

Fig. 4.16 shows the results of this experiment, plotted with the (positive abscissa) results for the previous experiment. As in the previous experiment, adding pixels along the figure boundary speeds search in a continuous, monotonic fashion. However, the addition of isolated, uniformly-spaced pixels has a much larger impact on search speed than the addition of pixels contiguous with existing contour fragments. The result is a highly nonlinear, concave response curve.

These results suggest a model which, rather than imposing tight constraints of connectedness on contour structure, not only allows the con-

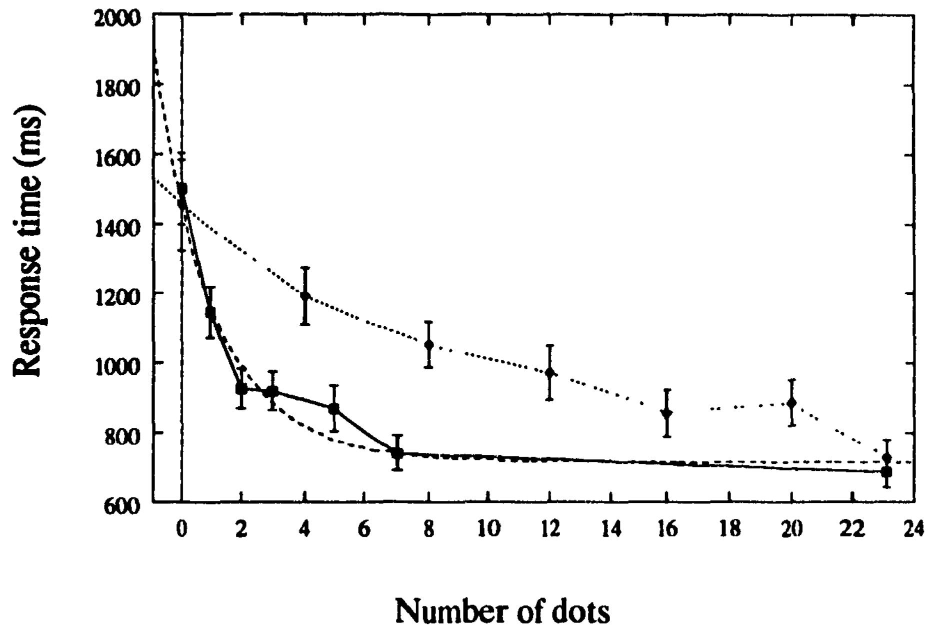


Figure 4.16: Results for boundary dot experiment (10 subjects, display size = 16). Results for continuous boundary extension experiment are shown dotted for reference.

tour information to be sparsely distributed, but in fact *prefers* it. Perhaps the most compelling evidence of this lies in the impact of adding just one pixel in the centre of the gap at both ends of the figure. In this case, we witness a very large improvement in search performance, while many contour grouping models would deny the addition of any oriented structure at all.

While a quadratic model for this data was found to be inappropriate, an exponential model (shown in Fig. 4.16) captures the shape of the data extremely well.[9]

However, there is a better way to look at these data. Rather than

[9]Each fit involved 630 data points (10 subjects × 9 trials × 7 stimulus conditions) The χ^2 for the quadratic model is 636, yielding $Q = 0$. The χ^2 for the exponential model is 627, yielding $Q = 1$.

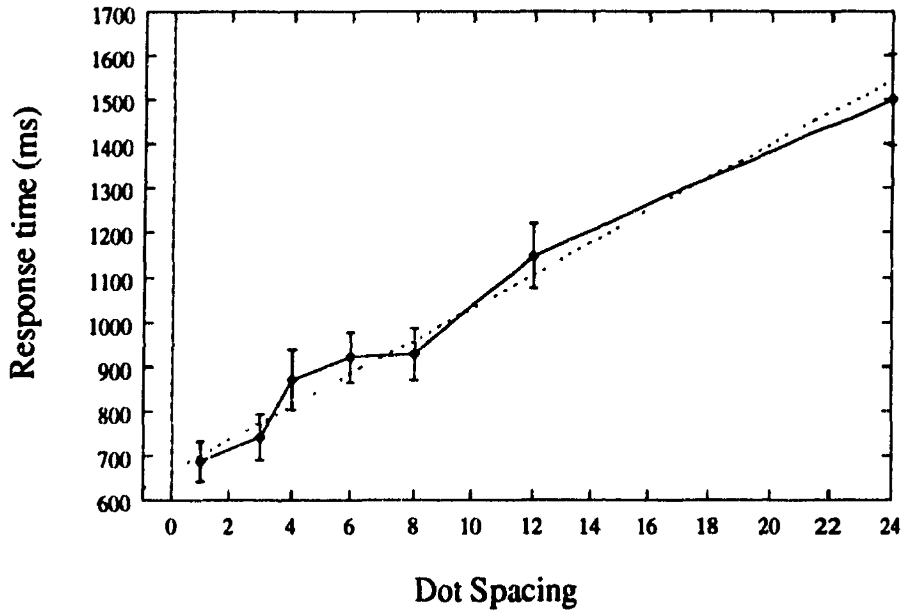


Figure 4.17: Linear fit to boundary dot search results, with response time plotted as a function of dot spacing.

treating the number of dots added as the independent variable, one can instead consider the *spacing* of dots in the gap. Fig. 4.17 shows that response time can be well-modeled as a simple linear function of dot spacing.[10]

Contrary to what one might expect, the value of contour in determining closure is not higher at points of high curvature. Note that this does *not* imply that support at points of high curvature is not important for determining the perception of local shape features. The degree of perceptual closure, however, is primarily a function of how well the contour is distributed about the boundary. To argue somewhat beyond the specific results of these experiments, I suggest that this last finding can be characterized as a **Minimax Gap Principle**: *Given a certain fraction*

[10]The χ^2 for this model is 625, yielding $Q = 1$.

of contour to bound a defined region, maximum closure is achieved by minimizing the maximum gap length.

The application of this principle depends upon the domain in which it is applied. In the continuous domain, the contour can be divided and distributed in such a way that the maximum gap is driven to zero. Perceptually, this would result in a closed contour with a reduced contrast. In a discrete domain, the uniformity of the distribution is limited by the pixel spacing, so that, given a contour which is shorter than the perimeter of the region to be bound, application of the Minimax Gap Principle will still produce finite boundary gaps.

The Minimax Gap Principle can also be stated in a manner less dependent upon the nature of the image space: *Given two different fragmentations of the same bounding contour, the fragmentation with the lesser maximum gap will have the greater closure.*

4.10 Textons are not Gluons

Consider the experimental results presented thus far. I have shown that, in a discrimination task requiring the representation of two-dimensional figure, performance is greatly determined by the degree of closure of the figure boundary. Properties such as connectedness, local corners and enclosure do not seem significant with respect to this effect. Psychophysical performance in this task varies as a continuous, monotonic function of closure, leading to the concept of a perceptual closure continuum. Greater closure is attained when contour is spread out around the figure boundary, rather than concentrated at points of high curvature. This observation has lead to the Minimax Principle stated above.

Most of these experiments have involved manipulation of the outline boundary of the figure: changing the boundary to increase closure seems to enhance the grouping of the contour fragments, sharpening the definition of figure and ground.

In most images, while there may be a high luminance gradient at the object boundary projections, there is typically some luminance pattern in the figure interiors which projects from reflectance variations over surfaces in the scene. Such patterns are typically referred to as *textures.*

Since in general different surfaces will project different texture patterns into an image, characterizing texture differences can potentially assist the segmentation of an image into regions or figures arising from distinct surfaces in a scene. This has motivated a broad program of research attempting to characterize and model the psychophysics of texture segmentation (Julesz [1991]; Malik & Perona [1990]).

Such region-based texture processes would seem to form a useful complement to the boundary-based processes I have been studying. In a real image, where boundary gradient signals are sometimes poor, region-based texture processes could potentially support figure/ground hypotheses.

The interaction of these boundary- and region-based processes can be studied in the context of my closure experiments. When closure is poor, there is ambiguity about whether the fragments should be interpreted as projections of a single object boundary. What if texture differences between the areas interior and exterior to the figures could resolve this ambiguity?

To answer this question, I have designed the three sets of textured stimuli shown in Fig. 4.18. In the middle set, I have augmented the original open stimuli with a texture of random dots, distributed over the two-dimensional region between the opposing contour segments. Each stimulus is thus defined by both contour fragments and a two-dimensional

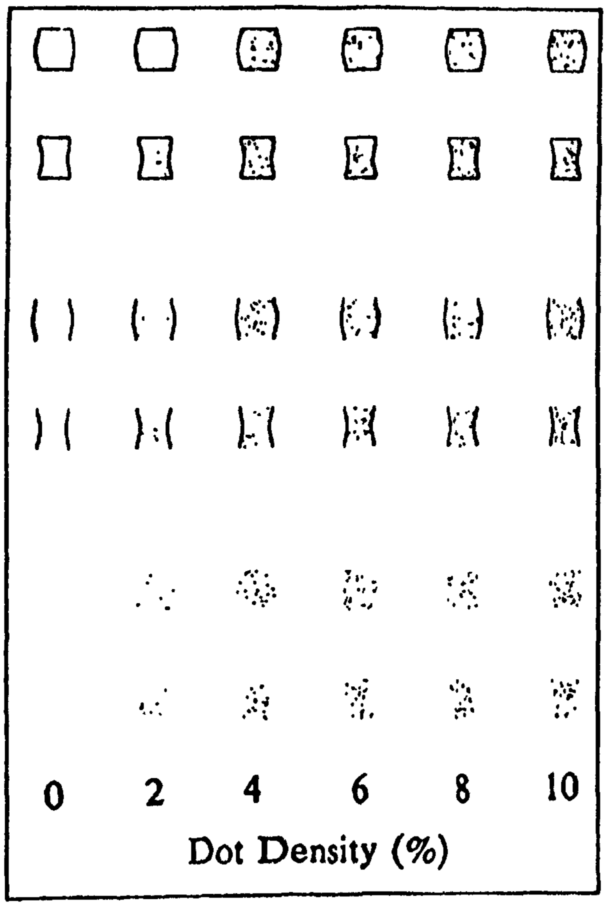


Figure 4.18: Stimuli used for texture experiment

texture, differing from the background in its first-order statistics (mean luminance). The independent variable in these stimuli is the interior dot density, which is varied between 0% and 10% in 2% increments. In the top set of stimuli, the same texture patterns are added to the original closed contours. In the bottom set, the stimuli are defined by the texture patterns alone. Note that although the texture used for each stimulus in a display is drawn pseudo-randomly from the same uniform distri-

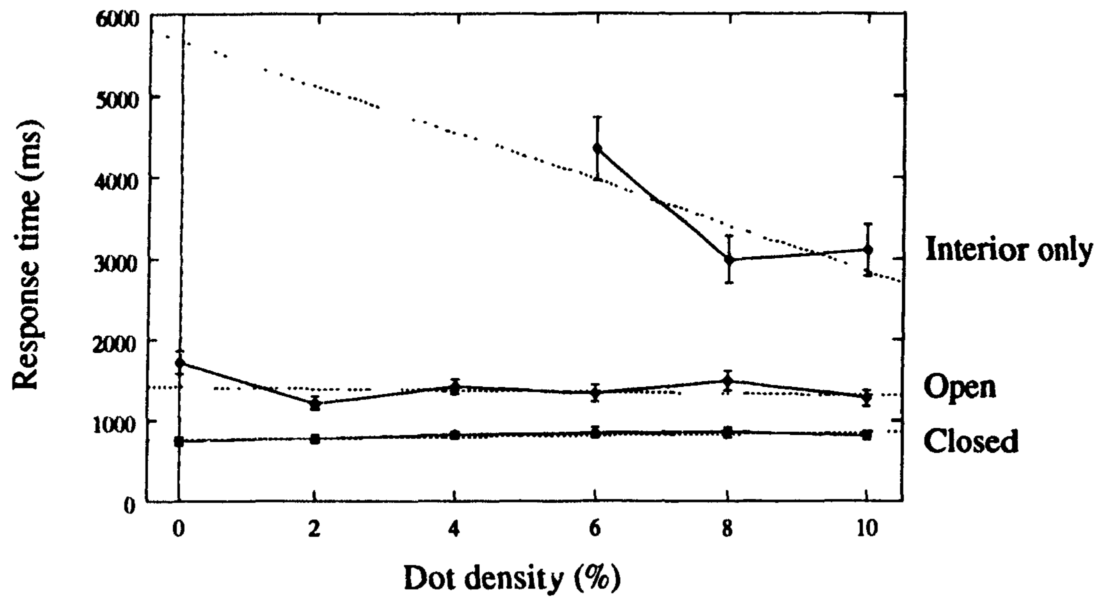


Figure 4.19: Search results for stimuli with interior texture (12 subjects). A display size of 16 was used.

bution, each is independent, resulting in a certain degree of distractor heterogeneity.

The results for these experiments are shown in Fig. 4.19. First, observe the response time curve for the stimuli with the interior dots alone. Data were not collected for densities less than 6% (about 40 dots per stimulus), because subjects typically could not discriminate these stimuli. Even for densities greater than 6%, the response time was very slow when no contour information was present. We can conclude from this that for the range of dot densities examined, the added dots do not make a significant *direct* contribution to the figure boundary. Thus any effect that we observe must arise from a region process serving to segment the textured figures from the background, thereby facilitating the grouping of the contour fragments into representations of two-dimensional shape.

Modeling the response time curves with linear functions suggests an insignificant effect of dot density on response time for both the closed and open contour stimuli: the gap in discrimination performance between open and closed stimuli is not narrowed by the addition of texture information.[11]

This is most remarkable in view of the dramatic sensitivity of search performance with respect to boundary manipulation. Whereas the addition of just 4 dots along the boundary of the figure cut the gap in search speed between closed and open stimuli by 70%, the addition of up to 75 dots to the figure interior has no significant effect.

4.11 Closure and Amodal Completion

Boundaries of occluded objects project as open and often fragmented contours in the image. I have argued that a figural system based upon a perceptual closure continuum will be robust to such occurrences, allowing the representation of object shape despite occlusion.

If robustness to occlusion is a prime determinant of the continuous character of perceptual closure, it is reasonable to expect that adding additional evidence for occlusion to the image would assist construction of shape representations and hence speed discrimination.

To test this possibility, I constructed the stimuli shown in Fig. 4.20. Starting with the partly closed figures on the left as a basis, I have created the set of figures shown in the middle by adding rectangle outlines to each partly closed figure to produce a subjective impression of occlusion. Note that in these figures, we perceive the spindle and barrel objects as

[11]Slopes of linear fits to response time as a function of texture density for open and closed figures do not differ significantly from 0 ($p > 0.1$) The intercepts (756 ms for closed figures, 1413 ms for open figures) do differ significantly ($p < 0.005$)

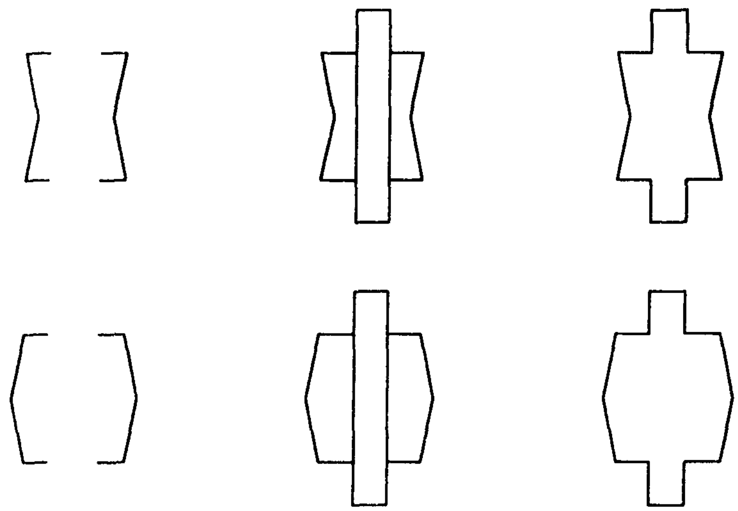


Figure 4.20: Stimuli used for occlusion experiment

being complete, but obscured by an intervening object. Kanizsa (Kanizsa [1979]) has termed this phenomenon *amodal completion*, because the perception of completeness is accomplished without direct verification by any sensory modality.

The set of figures on the right of Fig. 4.20 were used as control. These have been created from the second set by removing the contour segments of the occluding rectangle that fall in the interior of the original figure. These figures could be said to have the same 'shape' as the second set, but are seen as single closed figures with significant protrusions rather than as two figures in occlusion.

The results of these experiments (Fig. 4.21) are provocative in several ways. First, since the search for the occlusion figures (slope = 53 ms/item, intercept = 511 ms) is significantly *slower* than that for the original fragmented figures (slope = 27 ms/item, intercept = 535 ms)[12],

[12]Search slope for the occlusion figures is significantly greater than that for the fragmented figures ($p < 0.01$) Intercepts do not differ significantly ($p > 0.1$).

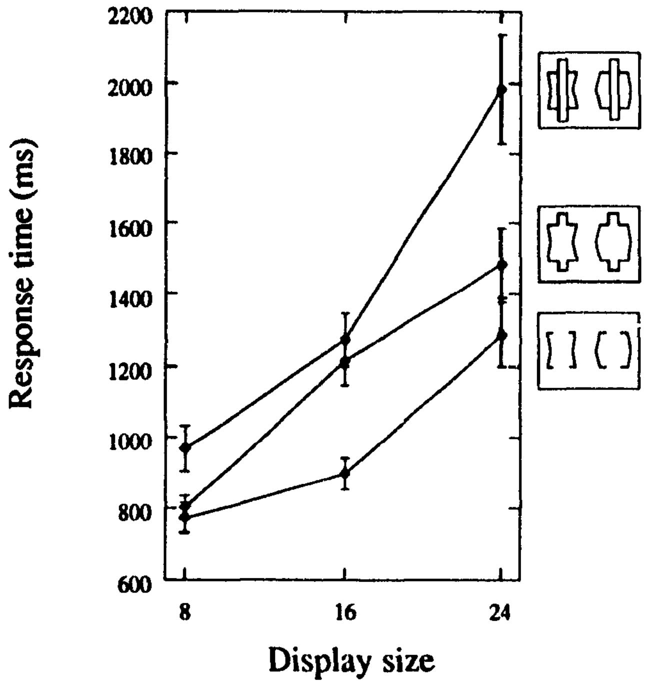


Figure 4.21: Search results for occlusion experiment (14 subjects).

the idea that introducing explicit occlusion cues would speed shape processing is clearly wrong.

The fact that search for the closed protrusion figures (slope = 45 ms/item, intercept = 454 ms) is also significantly slower than for the fragmented figures[13] suggests that the geometric complexity of the shapes has a strong impact on our ability to quickly discriminate them. We must be careful when we claim that closure leads to more rapid processing of two-dimensional shape: clearly the nature of the closing contour has a large effect on this processing.

[13]Search slope for the protrusion figures is significantly greater than that for the fragmented figures ($p < 0.01$). Intercepts do not differ significantly ($p > 0.1$)

4.12 Closure and Modal Completion

While boundary fragmentation can be caused by occlusion, it can also result from poor boundary contrast, caused by an unfortunate combination of photometric variables.

The response of the visual system to an extreme version of this problem is revealed by the phenomena known as *anomalous contours*. These are contours which, though defined in places by good luminance contrast, are without contrast support for significant stretches, and yet are perceived as uninterrupted. Kanizsa refers to these contours as *modally complete* because, even for the stretches without contrast support, the contours produce the *perception* of a luminance edge (Kanizsa [1979]).

While amodally completed contours are perceived as occluded, modally completed contours are perceived as *occluding*. In particular, most anomalous contour demonstrations involve basically regular silhouette figures with a sharp irregularity or incompleteness. This incompleteness can be explained if we suppose the figures to be occluded by another figure with reflectance equal to that of the background. There seem to be three components to this 'illusion': the perception of the anomalous contour itself, the perception that the occluding surface is brighter than the background, and the perception that it is closer than the occluded figures (Fig. 4.22).

The visual system thus appears to have a well-developed system for modally completing the contours of such figures. Could this completing process speed the formation of two-dimensional shape from fragmented image contours?

Fig. 4.23 shows stimuli designed to test this question. The stimuli on the left are modally complete. In the control stimuli shown in the

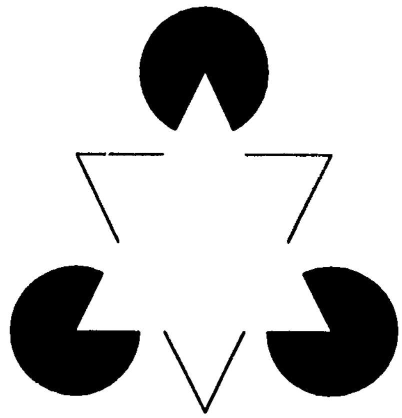


Figure 4.22: Classical example of modal completion

middle of the figure, the occluded silhouette figures are replaced by their outlines. On the right, the stimuli are further reduced to include only the components of the outlines lying along the occluding figure boundaries.

The results of these experiments (Fig. 4.24) fail to show any significant difference in search speed for modally completed contours (slope = 74 ms/item, intercept = 591 ms) over the fragment controls (slope = 55 ms/item, intercept = 388 ms).[14]

The perception of modally and amodally completed figures can be justified by the prevalence of occlusion and low boundary contrast in real images. In such cases, cues based on occlusion and figural incompleteness lead to the perception of whole figures despite boundary fragmentation. The fact that such clues do not speed shape discrimination reflects differences in *mechanisms* for low-level grouping based upon contour and more sophisticated processes requiring the inference of multiple overlapping surfaces. Note that to follow the 'logic' of modal or amodal completion one must first assert the existence of several overlapping surfaces.

[14]Neither slopes ($p > 0.05$) nor intercepts ($p < 0.1$) differed significantly

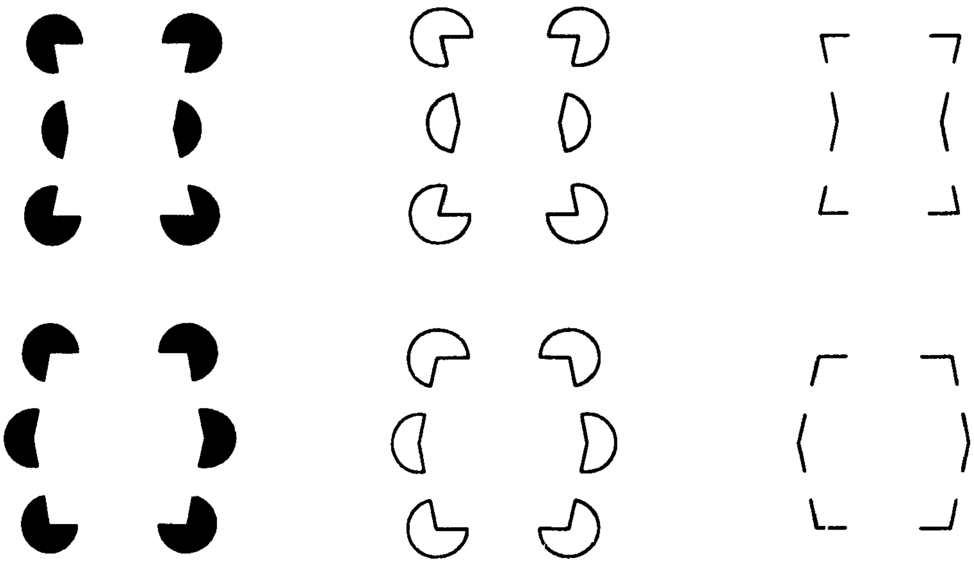


Figure 4.23: Stimuli used for modal completion experiment

It is likely that such assertions are predicated upon the very boundary processes that I have shown to be strongly determined by closure.

4.13 Search Asymmetry

In all of the experiments thus far described, the target has been spindle-shaped and the distractors have been barrel-shaped. This was an arbitrary choice, and it seemed worthwhile to measure search speed for the reverse experiment.

Fig. 4.25 shows the results of visual search experiments using the original open and closed stimuli of section 4.3 but with the roles of target and distractor reversed. The results show no asymmetry in search speed for the closed stimuli, but a significant asymmetry for the open stimuli: search is faster for an open barrel target in spindle distractors than for

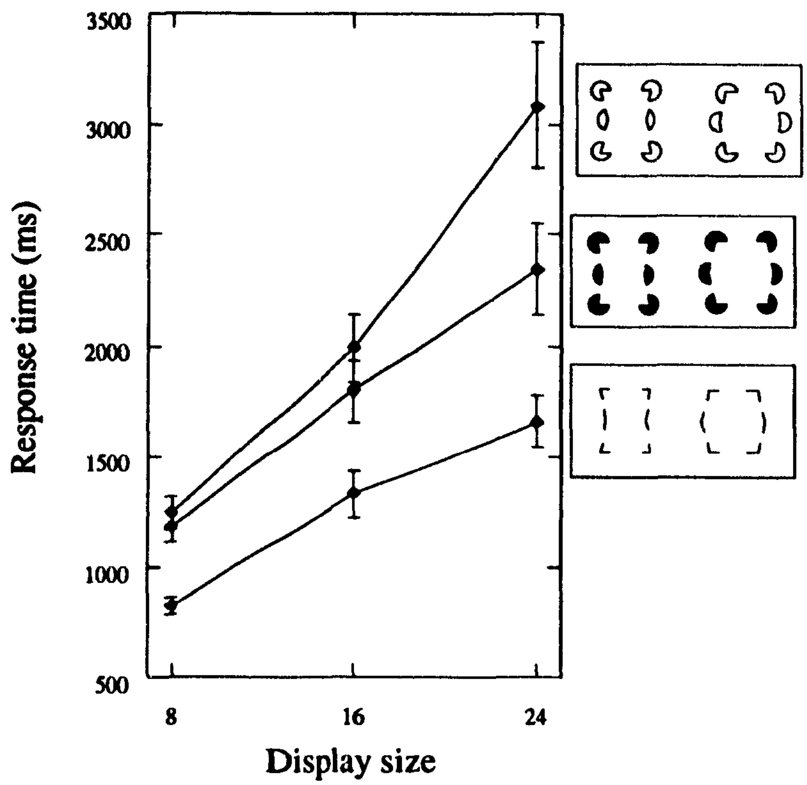


Figure 4.24: Search results for modal completion stimuli (14 subjects).

an open spindle target in barrel distractors.[15]

Treisman uses search asymmetry results to help identify perceptual features (Treisman & Gormican [1988]). However, the search asymmetry between open spindle and barrel figures admits a quite different interpretation. I believe that the asymmetry arises from the fact that the contour fragments making up the spindle figure are closer together than those for

---

[15]Search for a closed spindle target yielded a 14 ms/item slope and a 546 ms intercept. Search for a closed barrel target yielded a 15 ms/item slope and a 497 ms intercept. Neither slope nor intercept differ significantly ($p > 0.1$). Search for an open spindle target yielded an 83 ms/item slope and a 555 ms intercept. Search for an open barrel target yielded a 46 ms/item slope and a 677 ms intercept. Though intercepts do not differ significantly ($p > 0.1$), search slope for an open barrel target is significantly smaller than for an open spindle target ($p < 0.01$).

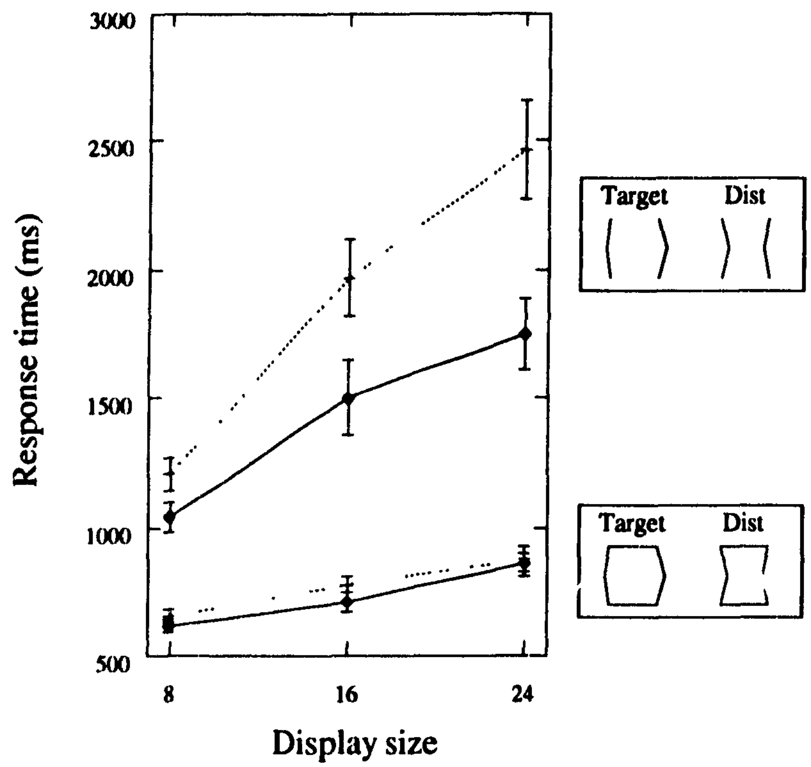
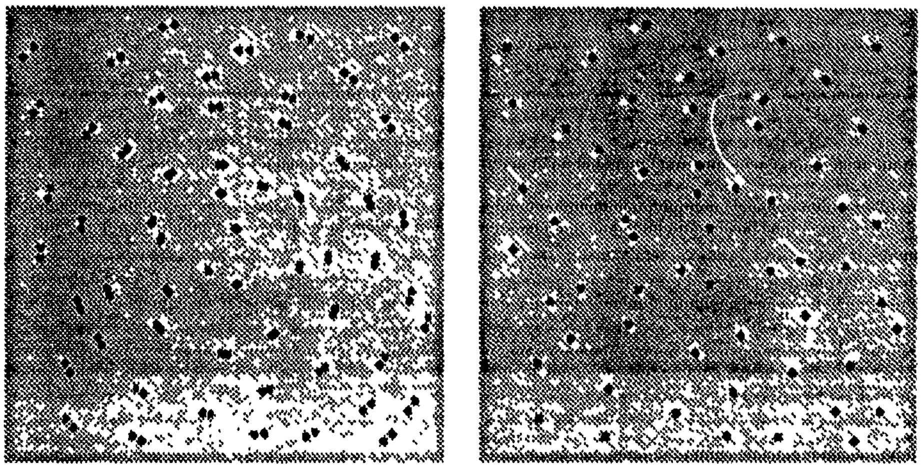


Figure 4.25: Search results for barrel target amongst spindle distractors (14 subjects). Results for spindle target amongst barrel distractors are shown dotted for comparison.

the barrel figure. This proximity advantage limits grouping ambiguity and leads to faster and more reliable figure representation for the open spindles than the open barrels. There will therefore be less ambiguity in a display with 23 open spindles and one open barrel than one with 23 open barrels and one open spindle. Less ambiguity leads to faster search on average.



(a) (b)

Figure 4.26: The destruction of short-range structure by contrast reversal. Pattern (a) is created by the superposition of a pattern of dots and a duplicate pattern rotated by 3 deg. Pattern (b) is identical to (a), with the duplicate pattern reversed in contrast. After (Glass & Switkes [1976]).

4.14 Closure and Contrast

From previous studies of contrast sign sensitivity, two classes of perceptual phenomena have emerged. Phenomena based on the short-range grouping of dots into one-dimensional structures (contours) can be destroyed by reversing the contrast of alternate dots (Glass & Switkes [1976]; Prazdny [1986]; Zucker & Davis [1988]; Zucker, Stevens & Sander [1983]). A beautiful example of this, due to Glass & Switkes, is shown in Fig. 4.26.

On the other hand, phenomena that depend upon longer-range grouping into two or three-dimensional structures seem to be robust to contrast reversal (Prazdny [1983]; Shapley & Gordon [1985]; Zucker [1986]). For example, (Prazdny [1983]) has shown that modal completion will occur

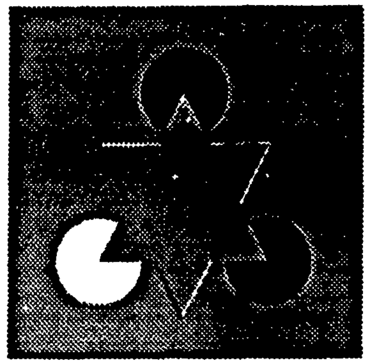


Figure 4.27: Modal completion of reversed-contrast contours.

Level	luminance (cd/m^2)
1	0
2	11
3	37
4	72

Table 4.1: Luminance values used in contrast experiments.

for contrast-reversing contours (Fig. 4.27).

If closure is a bridge between one- and two-dimensional structure, and between local measurements and global figural representation, into which class will it fall?

In experiments designed to answer this question, I used the four different luminance levels listed in table 4.1. Level 4, the brightest, is the luminance used to draw the figures in all of the experiments described to this point. Level 2 is the background luminance used in these experiments. Level 3 is an intermediate luminance which will be used to examine the effect of *reducing* figure contrast, and level 1 will be used to examine the effect of *reversing* contrast.

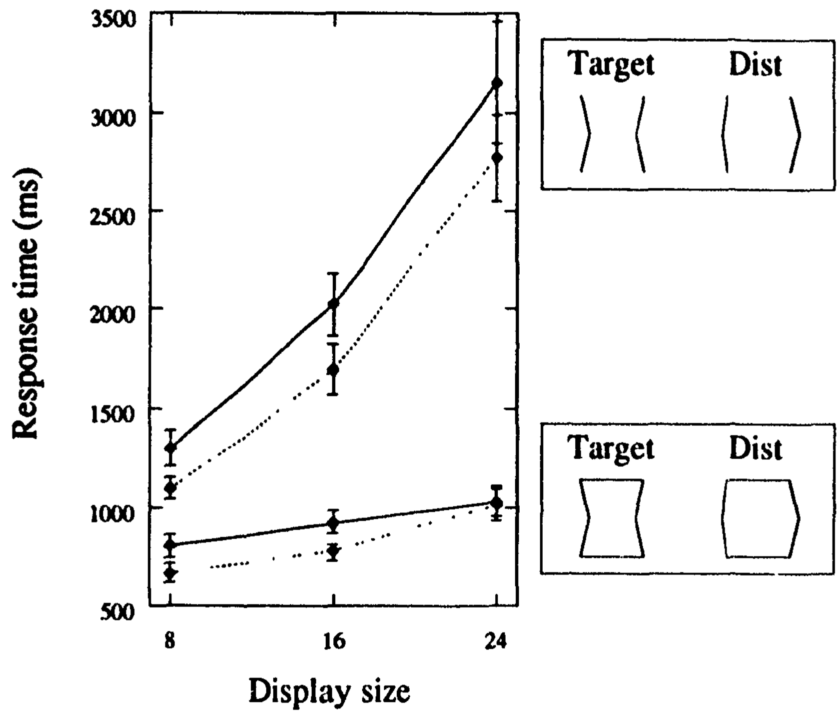


Figure 4.28: Search results for contrast control experiment (10 subjects). Results for black stimuli are shown solid, those for white stimuli are shown dotted.

4.14.1 Contrast Controls

The first experiment is a control to ensure that search speed does not depend upon whether the stimuli are drawn in white or black. The results (Fig. 4.28) fail to show any significant dependence of search speed upon the contrast sign of the stimuli.[16]

Our next two experiments are also control experiments, designed to determine whether variation in contrast over the entire display leads to

[16]Search using closed stimuli drawn in white yielded a 19 ms/item slope and a 496 ms intercept. Search using closed stimuli drawn in black yielded a search slope of 14 ms/item and an intercept of 692 ms. Search using open stimuli drawn in white yielded a 93 ms/item slope and a 340 ms intercept. Search using open stimuli drawn in black yielded a 105 ms/item slope and a 444 ms intercept. Stimulus contrast does not significantly affect search slope or intercept for closed or open stimuli ($p > 0.1$)

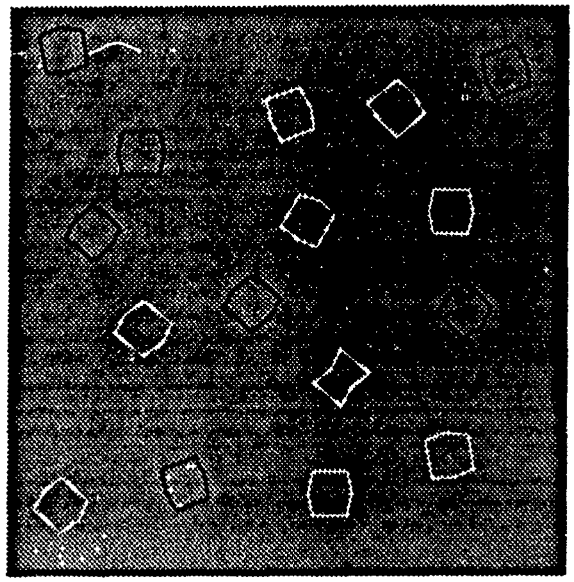


Figure 4.29: Search display with a mixture of black and white stimuli.

slower search speed. In both of these experiments, figures may be drawn either in black or in white (Fig. 4.29). In the first case, the target could be of either contrast. In the second case, the target was always white, and subjects were told this in advance. These experiments were performed only for the closed figures.

When subjects did not know the target contrast in advance (Fig. 4.30), search was slower than for the single contrast displays[17] (slope = 43 ms/item, intercept = 522 ms). However, when the target stimulus was always white, search was much faster (slope = 23 ms/item, intercept = 349 ms) and did not differ significantly from the single contrast search (Fig. 4.31).[18]

[17]Search slope is significantly greater than that for white stimuli ($p < 0.025$) and for black stimuli ($p < 0.01$). Intercepts do not differ significantly ($p > 0.1$).

[18]Search slope for the mixed stimulus displays where the target contrast is known does not differ significantly from slopes for all-white ($p > 0.1$) or all-black ($p > 0.05$) displays. Intercept for the mixed displays does not differ from that for the all-white

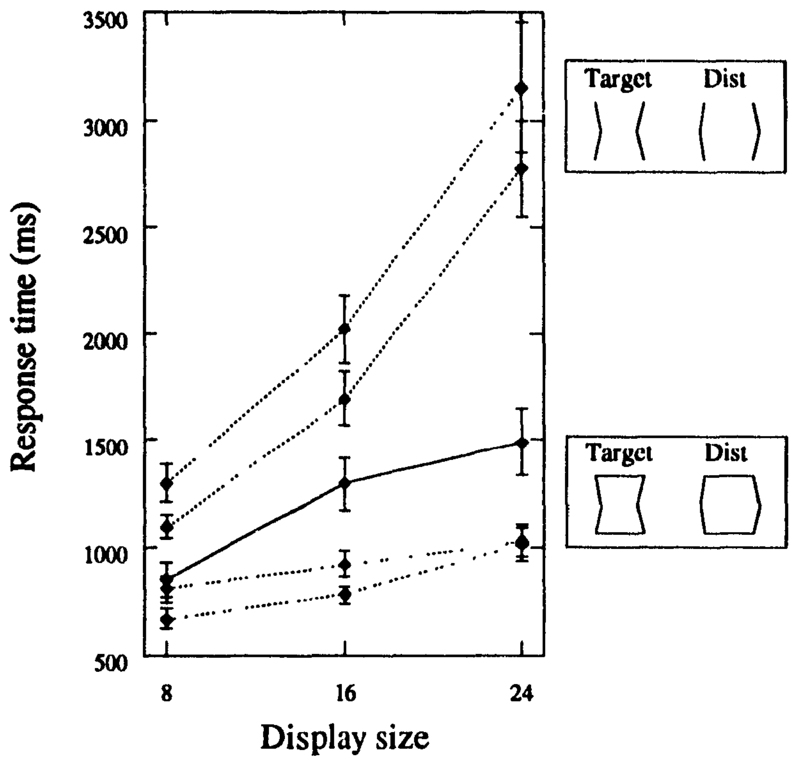


Figure 4.30: Results of search when subjects do not know the contrast sign of the target (10 subjects). Results for single-contrast search (either white or black) are shown dotted for reference.

I conclude from these two experiments that it is primarily the uncertainty in the nature of the target that leads to slower search speed: as long as the subject knows what to look for, contrast variation across the display does not significantly affect results.

4.14.2 Intra-Figure Contrast Variation

The control experiments have shown that search is independent of the contrast sign of the figures, and is not slowed by displays mixing figures

---

displays ($p > 0.05$), but is significantly *less* than that for the all-black displays ($p < 0.005$).

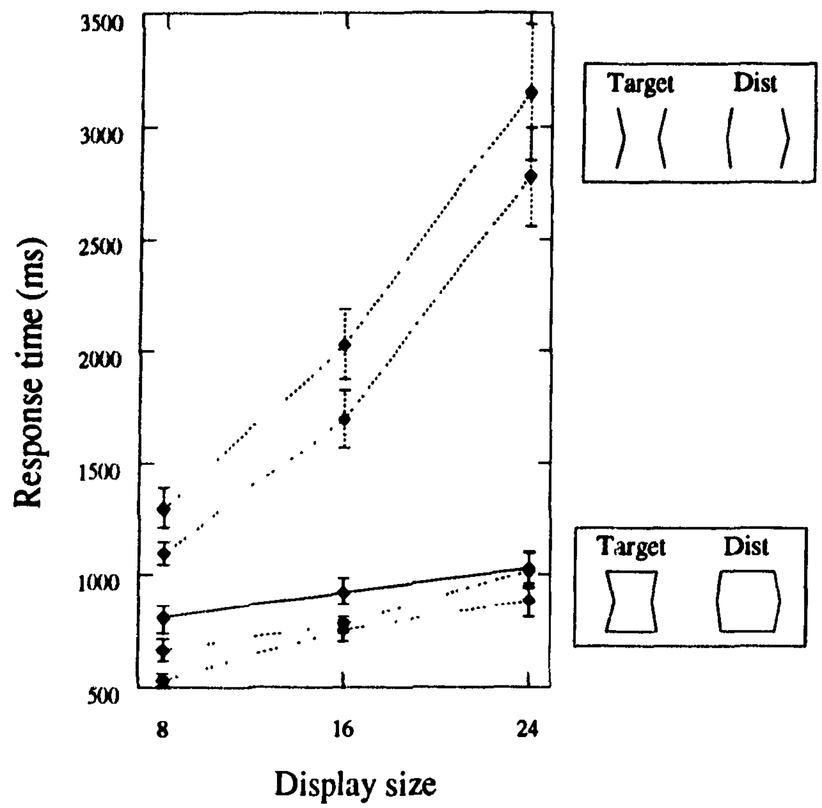


Figure 4.31: Results of search when subjects know the contrast sign of the target (10 subjects). Results for single-contrast search (either white or black) are shown dotted for reference.

of opposite contrast signs. We can now examine the effect of varying contrast along the contours of individual figures. The two stimulus pairs used for this experiment are shown in Fig. 4.32. In both pairs, the side fragments of the figures are drawn at luminance level 4. For the pair on the left, the connecting bars are drawn at luminance level 3, resulting in a *reduction* of contrast, whereas for the pair on the right, they are drawn at luminance level 1, resulting in a *reversal* of contrast. The results are shown in Fig. 4.33. While reducing the contrast results in only a mild decline in performance from the original closed figures (slope =

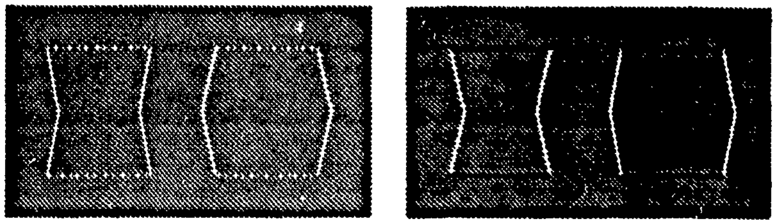


Figure 4.32: Closing the figures with *reduced* and *reversed* contrasts.

22 ms/item, intercept = 450 ms), reversing the contrast produces results nearly identical to those for the original open figures (slope = 90 m/item, intercept = 458 ms): contrast reversal eliminates perceptual closure.[19]

Repeating the reversal experiment with black figures and white closing bars (Fig. 4.34) produced similar results[20] (slope = 90 ms/item, intercept = 328 ms).

In these experiments contrast is reversed at the corners of the stimuli. Is it possible that search is slow because the visual system is unable to deal with a simultaneous discontinuity in orientation and contrast sign? To test this, I designed stimuli in which contrast reverses along straight sections of contour (Fig. 4.35).

Search speed for these figures is intermediate between that for the open and closed stimuli (Fig. 4.36).[21] This intermediate result reflects the fact that for these figures, integrating contour fragments only of the

[19]Reducing the contrast of the closing line segments produced a significantly greater search slope ($p < 0.05$), but had no significant effect on intercept. Reversing the contrast of the closing line segments resulted in mean search slope and intercept which do not differ significantly from those for the open stimuli ($p > 0.1$)

[20]Search slopes and intercepts for the open black stimuli and for the black stimuli closed with white contour segments did not differ significantly ($p > 0.1$)

[21]For the stimuli with short black connecting segments, slope = 48 ms/item, intercept = 473 ms. For the half-black, half-white stimuli, slope = 47 ms/item, intercept = 625 ms. Search slope for both experiments is significantly greater than that for closed stimuli and significantly less than that for open stimuli, of either contrast ($p < .005$) Intercepts do not differ significantly

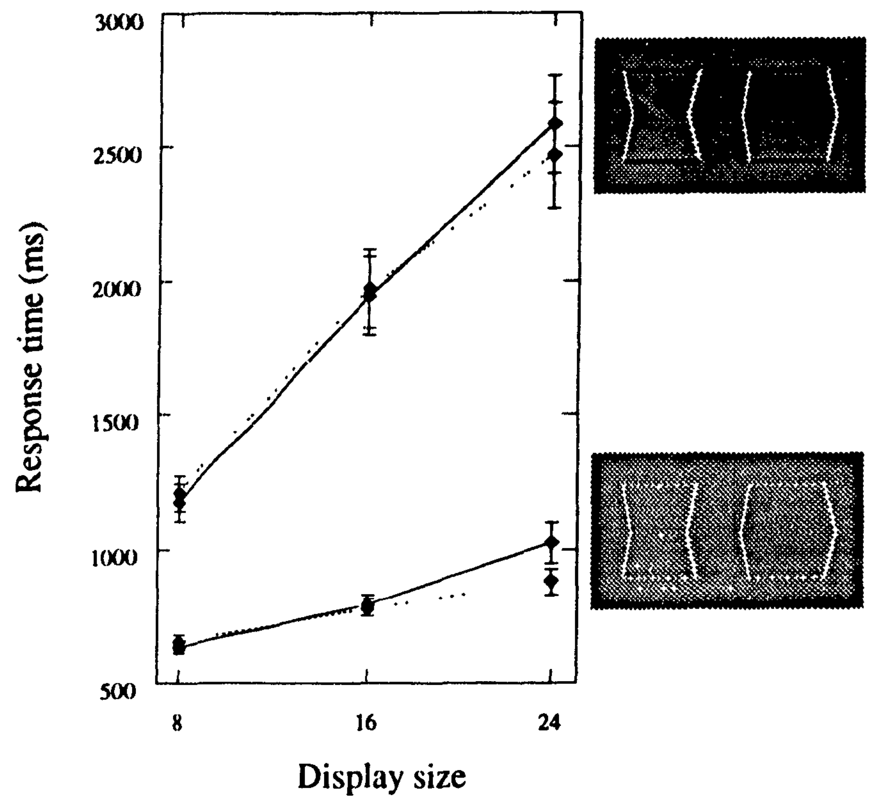


Figure 4.33: Search results for stimuli with reduced and reversed contrast fragments (10 subjects). Results for all-white stimuli are shown dotted for reference.

same contrast sign still yields an intermediate degree of closure.

The most important conclusion from these experiments I will call a Contrast Sign Principle: *perceptual closure operates only upon contour of a consistent contrast sign.* This result seems to conflict with recent models of shape processing which predict that boundary grouping processes must be insensitive to contrast sign (Grossberg & Mingolla [1985]; Shapley & Gordon [1985]). This prediction is based on the fact that contrast sign frequently *does* reverse along the boundary of an object which occludes distinct surfaces in a scene.

Relating my results to these theories is complicated by the fact that

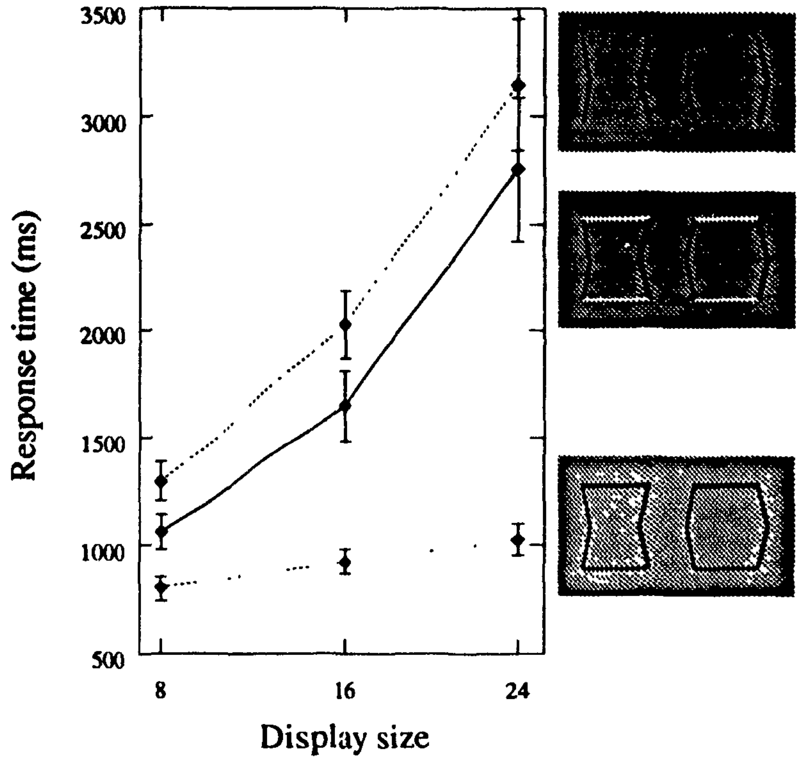


Figure 4.34: Search results for stimuli with black side segments and white connecting segments (14 subjects). Results for the all-black stimuli are shown dotted for reference.

my stimuli are drawn as outline figures. Typically we think of object boundaries as projecting luminance *edges*. Why then should we see figure just as easily from luminance lines?

To answer this question, we must first admit that the conventional model of an occlusion boundary projection as a step edge in an image is wrong. For a variety of reasons, sharp peaks and valleys of intensity are very common at and near occluding boundary projections (Perona & Malik [1990]). It is therefore possible that the visual system accepts and processes line drawings as approximations of occlusion contours, just as it would accept silhouette figures as such approximations.

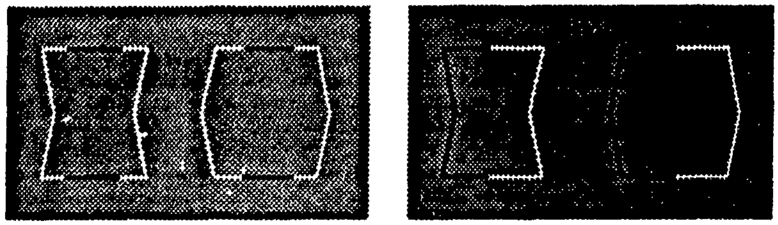


Figure 4.35: Reversing contrast along straight segments of contour.

In my experiments, sudden reversal in contrast is unsupported by accompanying shading changes on and around the occluding objects. Furthermore, because the stimuli are randomized in orientation, there is no simple lighting condition that could explain this luminance change. Faced with such an unnatural stimulus, it is not surprising that the visual system would not attempt to integrate contour of different contrast sign into single figures.

There is certainly adequate physiological support for a mechanism with such contrast-sign sensitivity. Experiments in cat visual cortex (Hammond & Mackay [1983]; Hammond & Mackay [1985]) show that not only are cells often contrast sign selective, they are highly nonlinear with respect to contrast sign reversal. Adding small amounts (e.g. 5% of total contour length) of reversed contrast contour can often completely inhibit a cell's response. This could provide a basis for fine-scale contrast sign-selective processing (Iverson & Zucker [1990])

The difference in contrast sign sensitivity between my experiments and phenomena such as modal completion (Fig. 4.27) reinforces the distinction between early shape-from-contour processes, and higher-level processes which organize the perception of multiple overlapping surfaces.

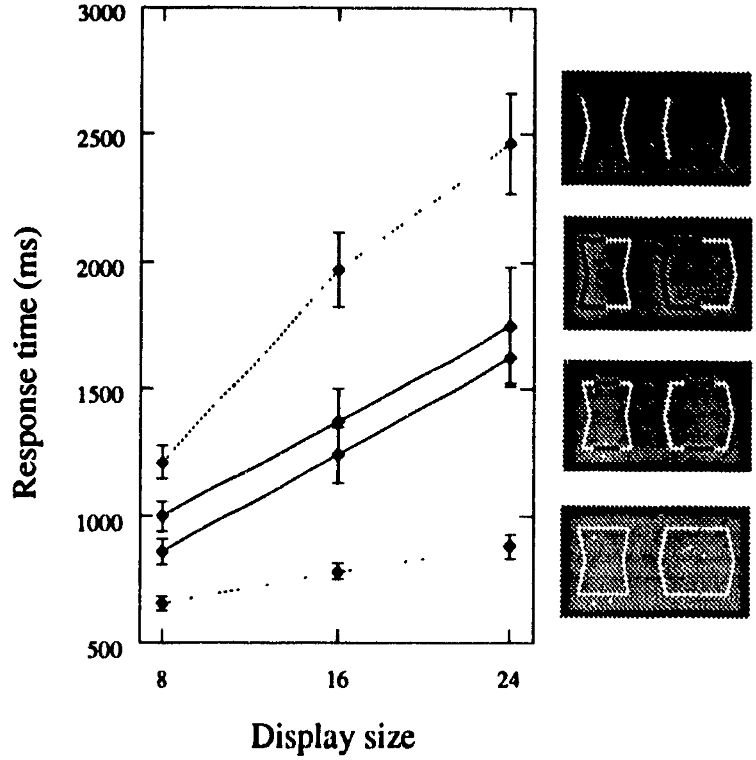


Figure 4.36: Search results for straight segment reversal (10 subjects). Results for the all-white stimuli are shown dotted for reference.

4.15 Discriminating Shape

The experiments discussed thus far demonstrate how contour closure determines our ability to discriminate two-dimensional shape, but do not identify the shape properties that are most important in this discrimination task. In this section, I will present two simple experiments which speak to this issue.

4.15.1 Size

One basic attribute of a shape is its size. In my experiments, the spindle figure is smaller than the barrel figure: in average width, maximum width

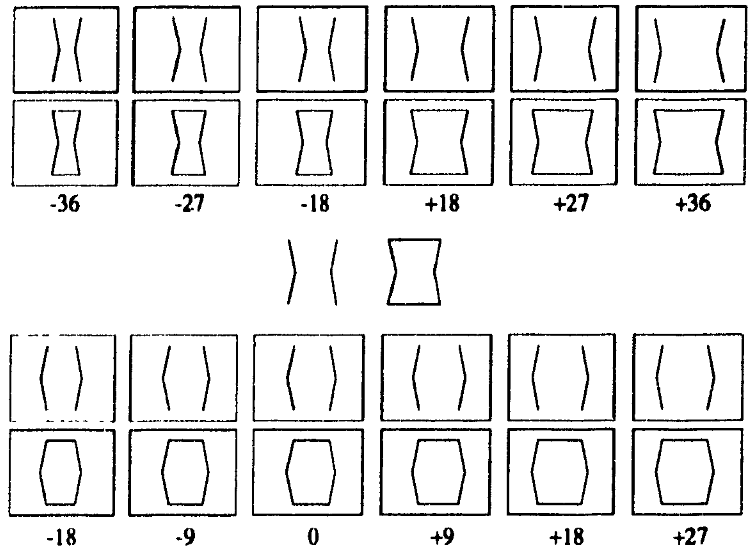


Figure 4.37: Stimuli used to test the importance of size in shape discrimination. The two figures in the centre are the original spindles, used as targets in this experiment. The percentage difference in size relative to these figures is shown beneath each of the distractors.

and area. To test whether size is the determining shape property in this discrimination task, I designed the stimuli shown in Fig. 4.37.

The two stimuli shown in the centre of the figure are the spindle figures used in my original closure experiment (section 4.3). They are used as targets in the present experiment.

The stimuli above these are spindles which have been uniformly narrowed or widened. The stimuli at the bottom of the figure are narrowed or widened barrels. The percentage difference in width (or equivalently, area) relative to the spindle targets is shown beneath each stimulus.

Four visual search experiments were conducted: two for open stimuli and two for closed. In each experiment, the target was a spindle of normal size. The distractors were either spindles or barrels of various sizes (Example displays for closed stimuli are shown in Fig. 4.38). The dis-

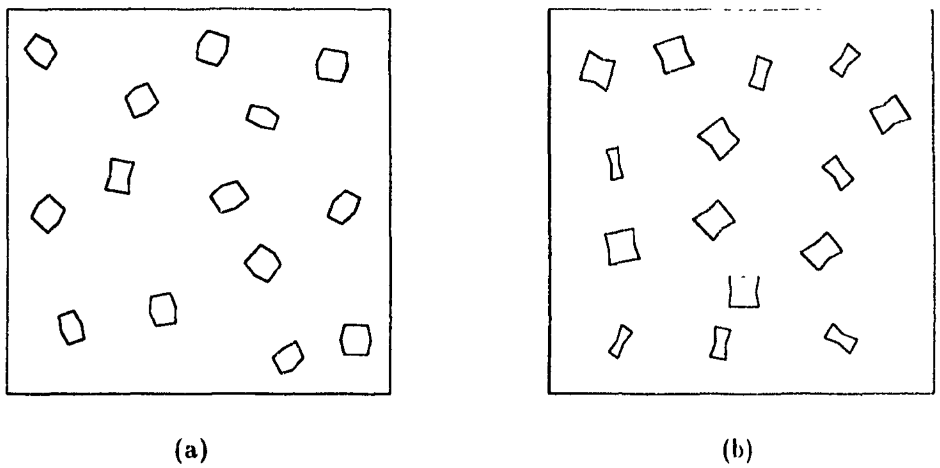


Figure 4.38: Example displays of visual search experiments based on difference in shape and difference in size.

tractors were chosen pseudo-randomly and uniformly from the stimulus sets shown in Fig. 4.37.

If size is the prime factor determining search speed, search should be faster for spindle distractors, which differ more from the target *in size* than do the barrel distractors.

The results are shown in Fig. 4.39. Search for the closed spindle target amongst closed barrels of different sizes (Slope = 26 ms/item, intercept = 470 ms) was only slightly slower than for the same experiment with barrels of constant size.[22] Search speed for the open spindle target amongst open barrels of different sizes (slope = 61 ms/item, intercept = 638 ms) does not differ significantly from the same experiment with barrels of constant size.[23] Search for the closed spindle target amongst

[22] Slope is significantly greater than for the closed figure experiment of section 4.3 ($p < 0.025$). Intercepts do not differ significantly ($p > 0.1$).

[23] Neither slope ($p > 0.05$) nor intercept ($p > 0.1$) differ significantly

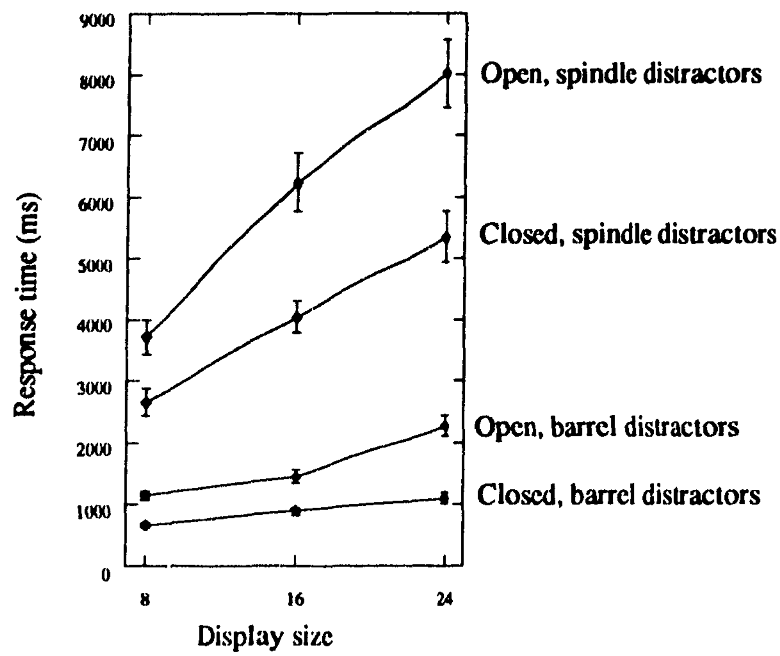


Figure 4.39: Results of experiments testing importance of size in shape discrimination (12 subjects).

spindle distractors (slope = 169 ms/item, intercept = 1312 ms) was much slower than the barrel distractor case.[24] Search for the open spindle target amongst spindle distractors (slope = 277 ms/item, intercept = 1545 ms) was also much slower than for the barrel distractor case.[25]

Using barrel distractors of various sizes prevented subjects from using size difference as a basis for discrimination, yet search was not appreciably slower than when barrels of a single size were used. On the other hand, when distractors were spindles either stretched or shrunk in width, search was very slow, indicating that size alone did not form a good basis for rapid discrimination. We must conclude from these results that size is

[24] Both slope ($p < 0.005$) and intercept ($p < .025$) were significantly smaller for barrel distractors than for spindle distractors.

[25] Both slope ($p < 0.005$) and intercept ($p < 0.05$) were significantly smaller when the distractors were barrels

not the prime determinant of shape discrimination for these figures.

4.15.2 Linear Filters

There have been several attempts to model visual search and texture discrimination with simple feed-forward systems based upon linear filters (Gurnsey & Browse [1989]; Julesz et al. [1973]; Malik & Perona [1990]). Typically, such models involve a convolution by local operators or 'detectors', followed by a more global measurement of changes in the statistics of the operator responses. In this section, I will show why these are not appropriate models for perceptual closure.

To be concrete, I will examine the texture model of (Malik & Perona [1990]). The algorithmic precision of this model allows a very direct analysis, which will make clear the gap between such models and the psychophysics of closure.

The model is based on a linear convolution with oriented and circularly-symmetric operators modelled on simple cells of striate cortex. Convolution is followed by two non-linear processing stages: a half-wave rectification and a spatially local lateral inhibition of operator responses. The final stage in this scheme involves the detection of large gradients in the response of these operators, which are then labelled as texture boundaries.

How could such a system lead to the discrimination of spindle from barrel? By virtue of their difference in shape and size, the response of the center-surround filters to the spindle and barrel will be different. This difference could be detected by the texture system and used to identify the target.

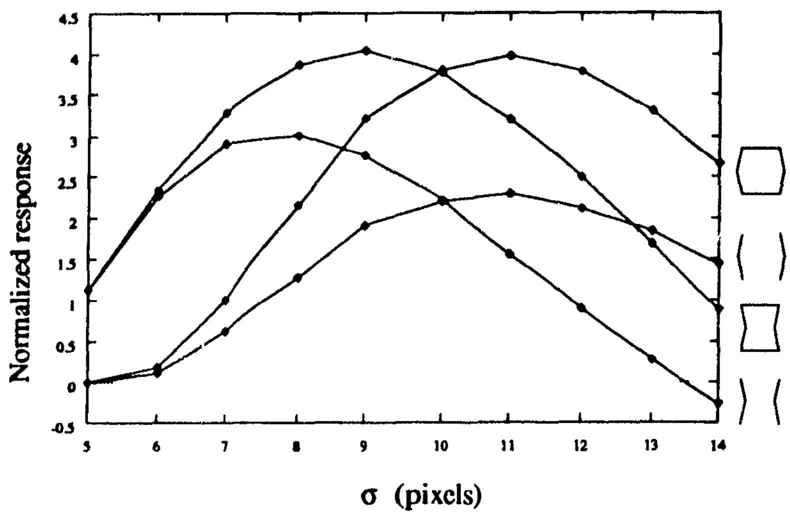


Figure 4.40: Filter responses to closed and open stimuli

The operators used are difference-of-Gaussian filters of the form

$$f(x,y) = \frac{1}{(0.71\sigma)^2}e^{-((\frac{x}{0.71\sigma})^2+(\frac{y}{0.71\sigma})^2)} - \frac{1}{(1.14\sigma)^2}e^{-((\frac{x}{1.14\sigma})^2+(\frac{y}{1.14\sigma})^2)}$$

.

The normalized filter responses for open and closed stimuli are shown in Fig. 4.40. Filter response peaks at a specific scale for each stimulus. The effect of the lateral inhibition stage will be to narrow these response curves around their peaks, suppressing inferior responses. Thus for the closed stimuli most energy will be concentrated in two channels. The spindle target is identified by a peak in the $\sigma = 9$ channel and a depression in the $\sigma = 11$ channel.

How could this theory account for the different search rates for open and closed figures? Notice that the responses to the open figures are 25-50% lower than the closed figure responses. If lateral inhibition weights were set appropriately (though it is not clear how this could be done *a priori*), oriented responses to the one-dimensional components of the stimuli, while suppressed by center-surround responses for the closed stimuli, might dominate in response to open stimuli. Since the target and distractor do not differ in their one-dimensional character, oriented filter responses do not differ either, and thus provide no means for discrimination. This could lead to slow search rates for the open stimuli.

Fig. 4.41 and Fig. 4.42 show the responses of the circularly symmetric filters to the closed spindle and barrel distractors of various sizes used in section 4.15.1. The response to the closed spindle target is also shown (solid line). Even after lateral inhibition, the energy is spread over many channels, excited by distractors of different sizes. Moreover, the channel which is optimal for the spindle target responds just as well or better for some of the barrel distractors. *There is no simple way to extract the target from this encoding.*

Of course, in the psychophysical experiments of section 4.15.1, subjects were able to discriminate the spindle from the barrels fairly rapidly, but were very slow to find the spindle target in different sized spindle distractors. There is no such distinction in the filter encoding shown in Fig. 4.41 and Fig. 4.42: if anything, the target is more unique amongst spindle distractors, since it is at least the only stimulus in the display with a peak response at $\sigma = 9$.

Of course, this linear model also does not capture the fine sensitivity of perceptual closure to small variations in the figure boundary such as the orientation of corners (section 4.4) or the addition of a single dot (section 4.9). A system based on channels of circle and line detectors tuned to different scales, and limited to the detection of gradients in the loudest channel, is too weak to model the psychophysical subtlety of perceptual closure.

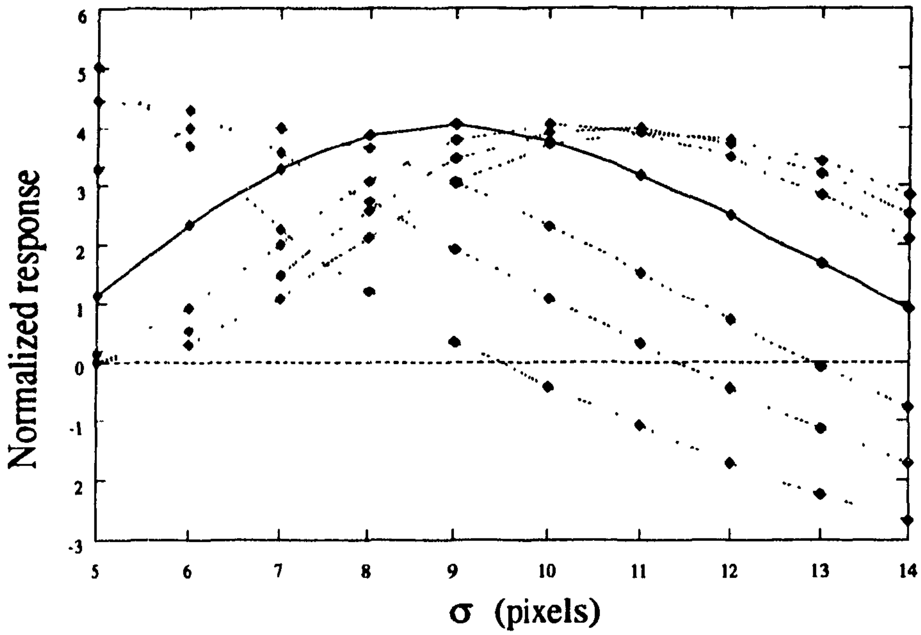


Figure 4.41: Filter responses to closed spindle target amongst spindle distractors. The target response is shown solid, the distractor responses are shown dotted.

4.15.3 Symmetry

The spindle and barrel figures both possess two axes of reflection symmetry. There is considerable evidence that symmetry plays an important role in perception (Biederman [1988]; Kanizsa [1979]; Rock [1983]). I was interested in how symmetry might interact with closure in determining the perception of two-dimensional shape.

To explore this question, I constructed the stimuli shown in Fig. 4.43. These figures were created by incrementally shearing the original stimuli to one side, allowing the contour to grow in length so that the height, width profile and area of the figures were maintained. The amount of shear (in pixels) is indicated for each stimulus.

I predicted that sheared figures would be harder to discriminate, for

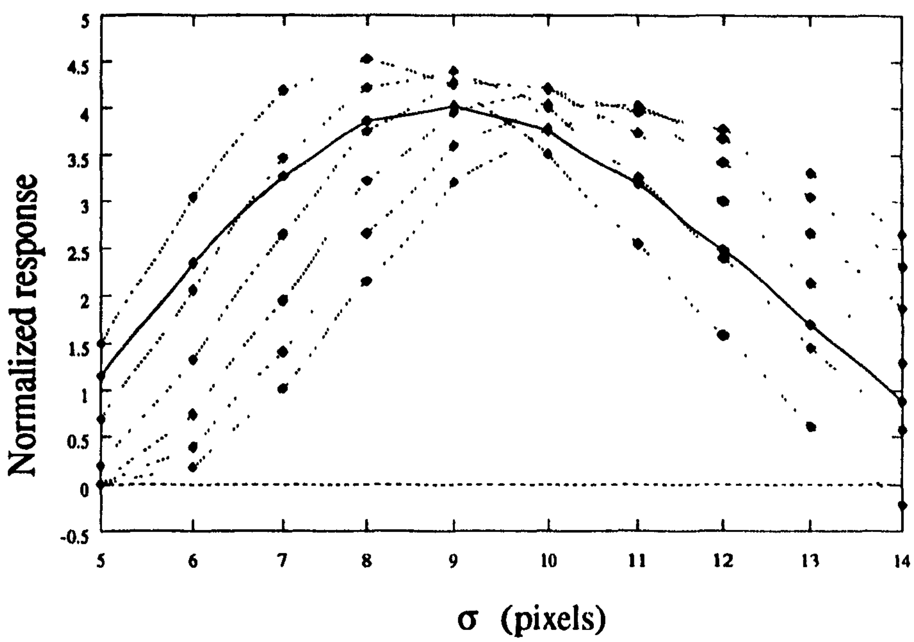


Figure 4.42: Filter responses to closed spindle target amongst barrel distractors. The target response is shown solid, the distractor responses are shown dotted.

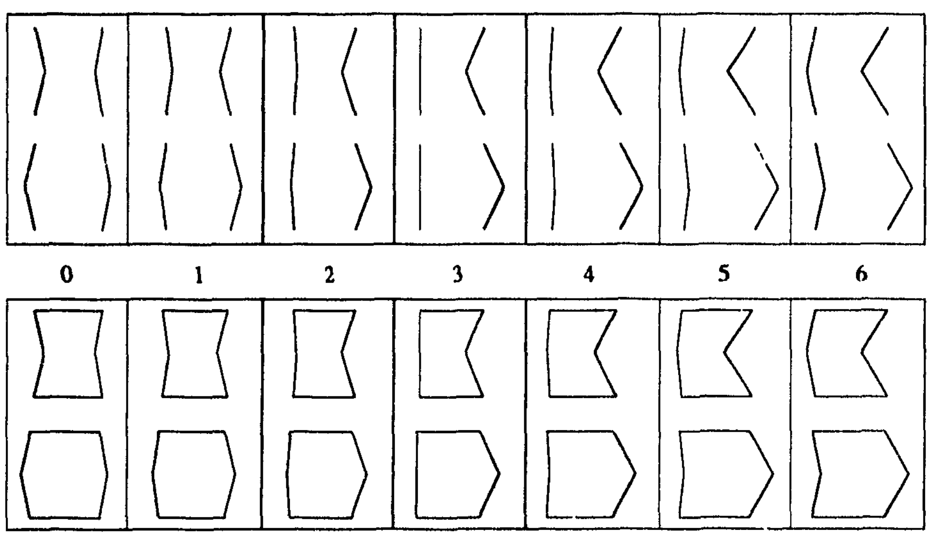


Figure 4.43: Sheared figures. The amount of shear (in pixels) is indicated for each stimulus.

two reasons. If symmetry is used in human perception as a kind of norm, sheared figures may lead to more complicated shape representations in which the information useful for discrimination is embedded. Accessing this information may thus be slower or less reliable.

Another consideration is that for shears greater than 3 pixels, the spindle loses a concavity and the barrel gains a concavity, so that each has one. If discrimination is based in part on such two-dimensional features, discrimination of the sheared shapes should be slower.

The results are shown in Fig. 4.44. While response time increases with shear for both closed and open shapes, the effect is about 6 times greater for open shapes.[26] I think this result is best interpreted by considering two views of symmetry: as both a grouping principle and as a shape quality. In the absence of good closure, the symmetry of the open shapes results in a closer association of the fragments and a faster or more reliable representation of two-dimensional shape. For the closed shapes, the discrimination is not slowed by grouping ambiguity and the effect is a reflection more of the shape representations than of their formation.

[26]For the closed shapes, slope = 48 ms/pixel shear, intercept = 612 ms For the open shapes, slope = 276 ms/pixel shear, intercept = 1380 ms.

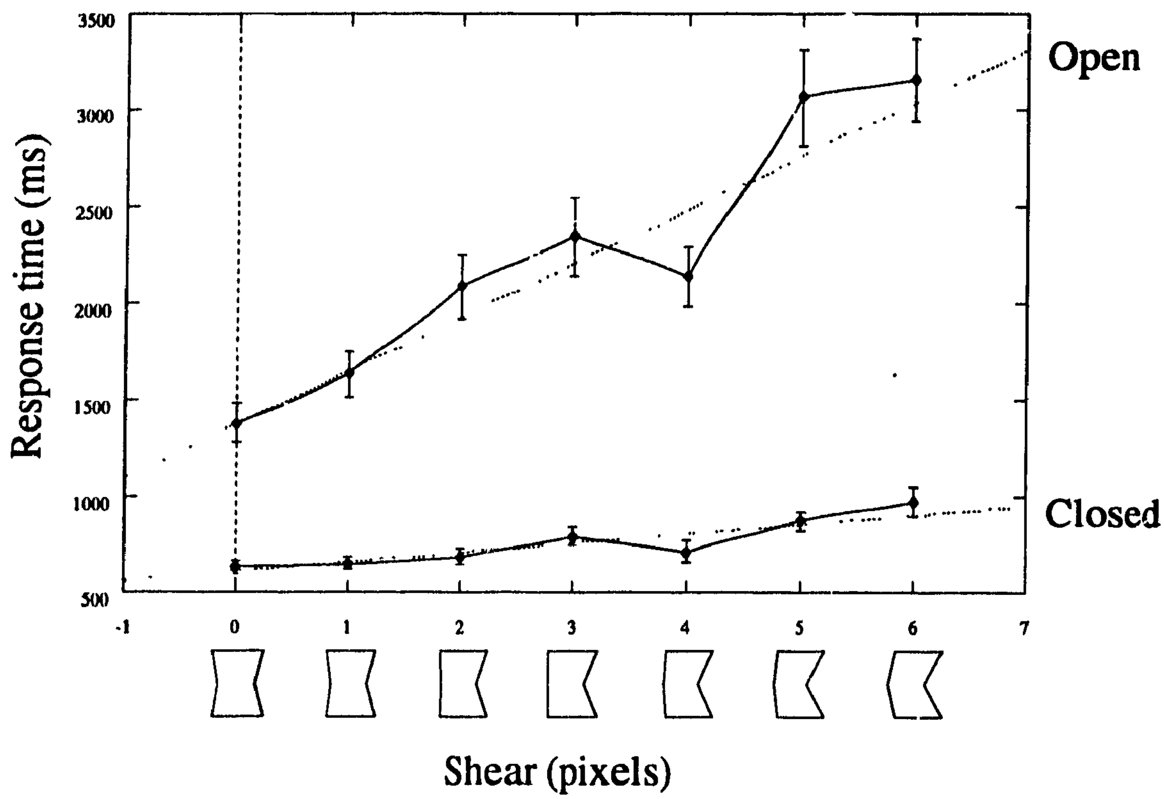


Figure 4.44: Search results for figure symmetry experiments (11 subjects). A display size of 16 was used.

Chapter 5 Discussion

5.1 Closure as a Measure of Single Object Confidence

A basic goal of vision is to integrate information from projections of objects into representations of object shape. Amongst the many contours in an image, there will be some which project from the boundaries of objects. Integration of contours from the boundary of an object allows the inference of two and three-dimensional shape properties which might be useful in a variety of tasks. The trick is to pick the right contours to integrate, so that non-occlusion contours, or occlusion contours from distinct objects, are not integrated into representations of non-existent objects.

The hypothesis motivating this work is that this process of selective integration is based upon contour closure. This perceptual closure is presumed to have some correspondence to mathematical and intuitive notions of closure, but also to have properties specific to a perceptual context.

My first experimental result confirms that two-dimensional shape discrimination is rapid for contours with good closure, and slow for contours with poor closure. This suggests that processing speed can be used (with care) as a measure of the *degree* of perceptual closure. Another useful way of viewing this metric is as a measure of the confidence with which contours can be interpreted as projections from a single object boundary.

A selection of the figures for which closure was measured in this way

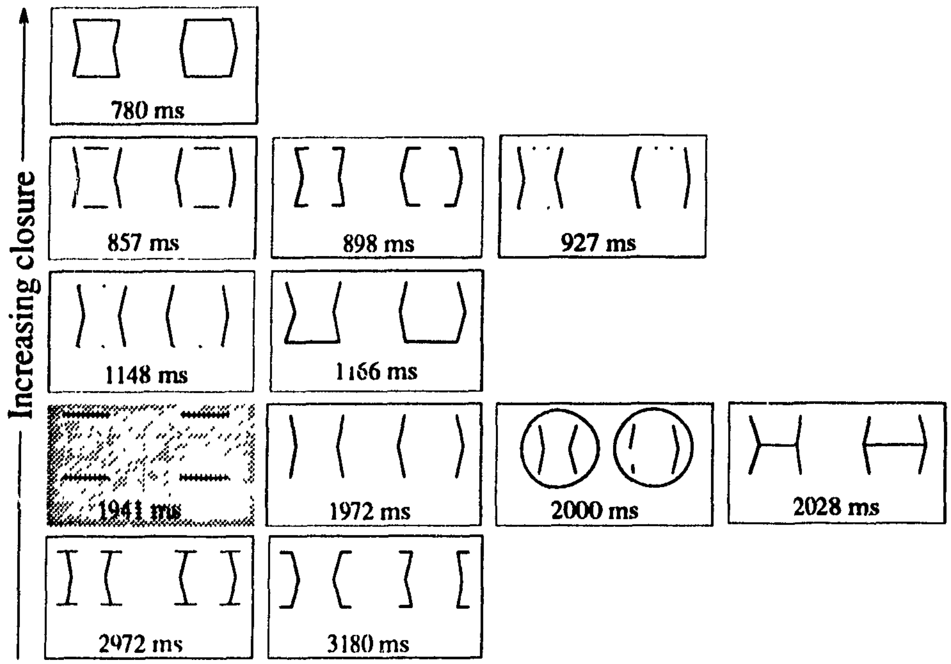


Figure 5.1: A summary of selected results. The figures have been ordered on the basis of the response time for a display size of 16. This response time is shown.

is shown in Fig. 5.1. Here I have ordered the figures on the basis of the mean response time for a display size of 16, so that the results for the stimuli with dots placed along the figure boundary could be compared with the other results.

Both geometric and photometric factors can cause an object boundary to project as an open or even disconnected (fragmented) image contour. Occlusion, changes in surface reflectance properties, and the occurrence or disappearance of specularities and cast or attached shadows can all contribute to this. On the other hand, these contour fragments could be interpreted as surface markings, or partial projections of multiple object boundaries. A visual system must consider these alternatives.

In section 4.8, search speed was measured for figures in which the contours were incrementally extended either along or away from the figure boundary. The result was a smooth, monotonic variation of search speed with contour extension, supporting the notion of a perceptual closure continuum.

This smooth relationship between closure and shape can be seen to balance two demands. While it is critical that perceptual processes be robust to contour fragmentation caused by occlusion or weak edge contrast, it is equally important that contour fragments arising from separate object boundaries or surface markings not be incorrectly grouped as projections of a single object boundary. As the degree of closure decreases, the "single-object confidence" decreases, and the plausibility of these latter interpretations increases. Since these interpretations would not suggest a computation of the properties which allow the shapes to be discriminated, dominance of these interpretations can be expected to slow the computation. The resolution of these ambiguities in favour of the single object boundary interpretation may require a greater contribution from reasoning processes.

In section 4.9, the dependence of response time on the distribution of contour around a figure boundary, and the observed linear relationship between response time and dot spacing led to the hypothesis of a **Minimax Gap Principle**: *given a length of contour and a region to be bound, maximum closure is obtained by minimizing maximum gap length.*

The physical justification for this principle lies in what has been aptly called *inferential leverage* (Witkin & Tenenbaum [1983]). The closure of the stimulus, that is, the confidence of the single-object interpretation, is closely related to the *conditional* probability that, given the observed image, the contour fragments project from a single object boundary. Two long but distant fragments could be interpreted as projections from two different objects. Fifty short contour segments, arranged evenly around a figure boundary to leave only small gaps, are unlikely to have arisen from distinct objects in the world.

In sections 4.4 and 4.6 I showed that properties of connectedness and enclosure, both suggested as principles of perceptual organization (Rock & Palmer [1990]) are not important in forming representations of shape from contour. Since contour fragmentation due to occlusion and photometric effects is so common, we cannot expect the boundary of an object to project as a connected contour in the image. Insensitivity to connectedness simply means that shape can be inferred from contour in the general case.

Although enclosure could be important for *preventing* grouping of an enclosed contour with a contour which is not enclosed, it does not increase the likelihood that enclosed contour fragments project from a single object boundary. Seen in this way, its lack of effect in determining shape discrimination is not surprising.

The experiments in section 4.14 establish the contrast sign sensitivity of perceptual closure: closure is computed only upon contour segments of the same contrast sign. If we assume line contour to be interpreted as an approximation or idealization of occlusion boundary projections (section 4.14.2) this sensitivity is justified by the absence of any plausible explanation of the luminance pattern in these displays based on lighting variation in a hypothetical scene.

The experiments of section 4.15.3 showed that the elimination of figure symmetry slows discrimination of the closed figures only mildly, but slows open figure discrimination dramatically. This suggests that, in the absence of closure information, the visual system exploits the low probability of two independent contour fragments projecting symmetrically.

The presence of symmetry between two contours in the image therefore supports the inference of a common cause: in this case a single object boundary.

5.2 The Why and How of Perceptual Closure

Many of my results (the Perceptual Closure Continuum, the Minimax Gap Principle, the Contrast Sign Principle, the effects of symmetry, the insensitivity to connectedness and enclosure) are consistent with the hypothesis that the speed of inference of shape from contour varies with the probability that the contour fragments project from a single object boundary. I have called this probability the *single object confidence*.

This hypothesis treats the visual system as an ideal machine, converting probabilities directly into performance: every aspect of the performance of the machine is based upon the statistics of scenes and images.

I have presented the perceptual closure continuum as a means for the human visual system to infer shape from contour despite occlusion and low reflectance contrast. Introducing explicit evidence of occlusion or low reflectance contrast by amodally or modally completing the figures should increase the single object confidence by providing "explanations" for the contour fragmentation. The single object confidence hypothesis therefore predicts that the perception of shape from contour should be speeded when the contours are amodally or modally completed.

The results presented in sections 4.11 and 4.12 do not support this prediction: amodally completing the figures slows search significantly, and modally completing the contours shows no significant effect. This seems like a contradiction: I am claiming that the purpose of the closure continuum is to allow the rapid inference of shape from contour fragmented by occlusion or low edge contrast, but when explicit cues to these effects are added to the displays, discrimination is slowed.

This apparent contradiction is resolved by a careful distinction between the *why* and the *how* of perception. The advantages of being able to perceive shape despite occlusion or shadowing may have driven the evolution of a perceptual closure continuum, and thus account for our ability to infer shape from fragmented contour. However, this does not mean that shape will necessarily be perceived more quickly if explicit occlusion cues are introduced.

While theories of the *why* of perception need not consider the computational complexity of a task, theories of the *how* must. Overlaying the partly closed figures with a rectangle outline (Fig. 4.20) gives the visual system a lot more to compute: the existence and shape of multiple surfaces overlapping in depth. Discrimination of the occluded shapes is made more difficult because there is more coded data to sort through: we are not able to completely ignore irrelevant information.

The fact that shape can be rapidly inferred from fragmented contour without explicit occlusion cues supports the existence of a limited-complexity shape-from-contour system which is able to function independently of more complex mechanism required to process occluding surfaces.

A similar result was obtained in section 4.10. It seems reasonable to believe that a difference in texture between figures and background would increase the single object confidence: it is unlikely that random dots would by chance only lie between two proximal contour segments, unless the surfaces between the proximal fragments were distinct from the background surface. Given the interpretation of the contour fragments as occlusion boundary projections, the interpretation of the dots as elements of surface texture is plausible, yet it does not assist the performance of the task.

This result further supports the existence of an early shape-from-contour boundary process which does not interact with region processes at an early stage.

5.3 Mathematical Tools for Closure

It would be nice if we could formalize the notion of perceptual closure into a tight mathematical statement. Unfortunately, the fact that perceptual closure is so independent of the property of connectedness makes it unlikely that such a simple theory could be formed. In this section, I review several mathematical ideas relevant to the topic of contour and demonstrate the difficulties in extending these tools to perception.

5.3.1 The Jordan Curve Theorem

A Jordan curve is defined as any figure which is topologically equivalent to a circle. The Jordan curve theorem can be formally stated as:

> Let $\mathcal{J}$ be a Jordan curve. Then the complement of $\mathcal{J}$ in the plane, $\mathcal{J}'$, is not connected but consists of two disjoint connected pieces, one of which is bounded (called the **inside**) and one of which is not bounded (called the **outside**). The curve $\mathcal{J}$ forms the boundary for both pieces (Henle [1979]).

Since the distinction of inside and outside allows us to unambiguously define curvature sign, this theorem is essential to computational theories of shape from contour. Constraints on surface shape derived from the curvature of an occluding contour (Biederman [1988]; Koenderink [1984];

Lowe [1985]; Marr [1982]) require that the sign of curvature be defined. Curvature sign is also required for many computational theories of planar shape (Blum [1973]; Hoffman & Richards [1985]; Kimia, Tannenbaum & Zucker [1990]; Leyton [1989]).

The problem is that ***topological connectedness is not a stable property of image contours.*** This is just another way of saying that occlusion and poor image contrast can lead to contour fragmentation.

Perceptual closure is thus not simply a topological property: any model will at least require geometric tools as well. I will briefly mention two.

5.3.2 The Theorem of Turning Tangents

Geometrically, a regular plane curve $\alpha : [a,b] \rightarrow R^2$ is closed if α and all its derivatives agree at a and b. The curve α is simple if it has no further self-intersections. If $\alpha(s)$ is a regular, planar, unit speed, C^2 curve we can define $\theta = \int_a^b k(s)ds$, where $k(s)$ is the curvature of $\alpha(s)$, with sign chosen arbitrarily. A theorem often called the Theorem of Turning Tangents (do Carmo [1976]) provides that if α is simple and closed, $\theta(s) = \pm 2\pi$.

This theorem could provide a closure metric for open curves: values of θ near $\pm 2\pi$ indicate contours with a high degree of closure, small values of θ indicate a low degree of closure.

Although this metric is defined only on C^2 curves, we could generalize it to piecewise C^2 curves by defining $\theta(s)$ at points of curvature discontinuity to be the signed difference of nearby tangents.

Generalizing this metric to a piecewise continuous curve is not easy. First, the metric will depend heavily on how we choose the sign of cur-

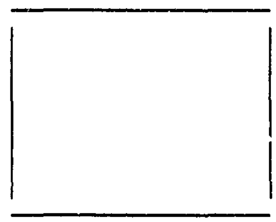


Figure 5.2: Figure with 0 closure using a curvature metric.

vature for each contour fragment. Even if we found some principled way to do this (using proximity of endpoints, for example), there are still many fragmented contours with a high degree of perceptual closure which would have value 0 under this metric (Fig. 5.2).

5.3.3 The Winding Number of a Curve

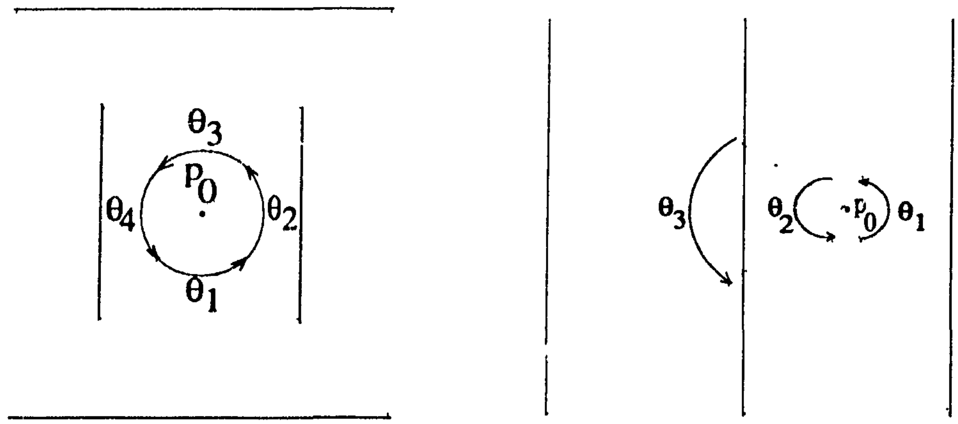
Let $\alpha : [a,b] \to R^2$ be a closed planar curve. Choose a point $p_0 \in R^2$, $p_0 \notin \alpha([a,b])$, and let $\phi : [a,b] \to S^1$ be given by

$$\phi(t) = \frac{\alpha(t) - p_0}{|\alpha(t) - p_0|}, \qquad t \in [a,b].$$

We can (informally) define the *degree* of ϕ with respect to p_0 as the number of times that $\phi : [a,b] \to S^1$ wraps $[a,b]$ around S^1. The degree of ϕ is called the *winding number* of the curve α relative to p_0.

It can be shown that if two points in the plane can be connected by a path which does not intersect the curve α, then α has the same winding number relative to both of these two points. In particular, the winding number of a simple, closed curve relative to a point in its interior is 1, and relative to a point in its exterior is 0.

We could generalize the definition of winding number to apply to continuous curves which are not closed. However, with this generalization we



(a) (b)

Figure 5.3: Open figures with large winding numbers.

lose the stability of the winding number with respect to p_0. Further extending the winding number to piecewise continuous (fragmented) curves would encounter the same problem as for the Theorem of Turning Tangents: dependence upon the parameterization of each continuous contour segment. Although such an extension might provide a more useful definition of closure for contours such as that shown in Fig. 5.2, there are other examples for which the metric would again provide very unintuitive answers (Fig. 5.3).

The Jordan Curve Theorem, although it may provide little operational guidance, is important as a formal statement connecting contour closure to two-dimensional shape. Geometric views on contour closure also embody notions of importance. The Theorem of Turning Tangents provides an intrinsic view connecting contour curvature and contour closure (a curve has a high degree of closure if it is, on average, curving significantly to one side), and the concepts of degree and winding number provide an extrinsic view based upon the enclosure of a point.

In Fig. 5.3(a), while the chosen p_0 results in a winding number near 1, other choices yield a much smaller measure. This suggests that any extrinsic computation of closure cannot be so point-dependent, but rather should be based upon an integration over the the entire figure interior. Fig. 5.3(b) shows that the relative positioning of fragments cannot be ignored.

While development of a formalism which captures these concerns is beyond the scope of this work, a short metaphorical aside may provide some inspiration.

5.4 Metaphor

Imagine an image as a three-dimensional surface, where contours are represented as walls, and the space between contours as flat planes below these walls. Imagine a divine intervenor with a sweet tooth who wields a large bucket of honey and begins to pour it over this surface. Due to its surface tension, the honey will not simply flatten out uniformly, but rather will tend to be enclosed within certain regions (Fig. 5.4). What if we define the closure of contour fragments as the height of the honey enclosed?

This metaphor has a certain appeal. First of all, it captures the notion that boundaries are inherently enclosers: they contain *stuff*. Secondly, with the introduction of surface tension, the metaphor allows for a continuous definition of contour closure over fragmented curves. Even better, the maximum gap principle finds a home, since the total resistive force of many regularly-spaced ridges will be greater than a more continuous ridge with fewer but larger gaps.

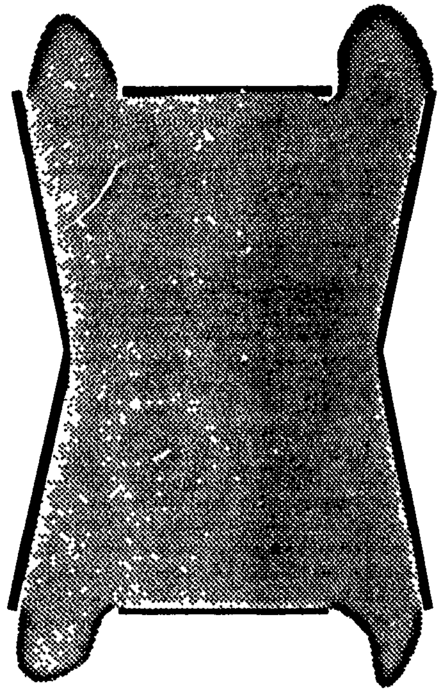


Figure 5.4: Closure as a measure of the ability to contain *stuff.*

Limitations of the metaphor can also be expressed in terms of my experiments. Clearly outward-pointing corners have no influence on the ability of two contour fragments to contain honey, yet we know from the experiments of section 4.4 that such corners have a large influence on perceptual closure. This example suggests that perceptual closure cannot be described by an extrinsic model alone: local boundary geometry must be considered.

5.5 The Nonlinear Nature of Perceptual Closure

Linear transforms are very popular in computer vision: consider for example the scale-space techniques used to describe the properties of a signal at a variety of spatial frequency subbands (Burt & Adelson [1983];

Witkin [1983]; Witkin, Terzopoulos & Kass [1987]). In this section I will show that perceptual closure is inherently *non*linear, and will provide an example of a nonlinear mechanisms which could model one aspect of my experimental results.

The Contrast Sign Principle, derived from the experiments of section 4.14, states that perceptual closure is a function only of contour of consistent contrast sign. In other words, perceptual closure simply ignores contour which is not of the right contrast sign. This behaviour is highly nonlinear: contrast reversal would have an *inhibiting* effect were the system linear.

The sensitivity to the orientation of corners is disproportionate to the fraction of contour which they represent (section 4.4). The effects of dots added in the boundary gaps of the open figures is also disproportionate: adding a single dot in the middle of the gaps of the open figures cuts the difference in search speed between closed and open figures by about 44%. Adding two dots cuts it by 70% (section 4.9).

These boundary dot experiments show that perceptual closure is not only a highly nonlinear function of the *amount* of contour added, but that it is also very dependent upon *where* contour is added *relative* to where contour already exists. The effect of an added input to the system depends upon the existing input: the principle of superposition is violated, and the system is therefore nonlinear.

This violation is also reflected in the difference between the perceptual closure of a stimulus which is closed at one end and open at the other and a stimulus which is half-closed at both ends (section 4.4). Consistent with the Minimax Principle, the figure which is half-closed at both ends has a greater perceptual closure, though the total boundary gap is the same for both stimuli.

A further example of the violation of superposition can be seen in the experiment of section 4.5. While adding inward corners to the open stimuli *speeds* discrimination, adding inward corners to the stimuli with outward corners *slows* discrimination.

The Minimax Principle states that ***maximal closure is attained by minimizing the maximum boundary gap.*** One type of nonlinear processing that could lead to the Minimax Principle is shown in Fig. 5.5. The figure depicts a series of colinear dots in an image. These dots are integrated by operators with small receptive fields (represented by the small boxes). The response of each operator is a function of the number of dots within its receptive field. This function is nonlinear, with a negative curvature. The responses of these operators are integrated by a larger-scale operator.

The negative curvature of the response functions means that the incremental response to dots added within an operator's receptive field will diminish with each dot added. In order to maximize the total response, dots must be spread out so that the response of each operator stays on the steep portion of its response function. In other words, a potential basis for the Minimax Principle of Closure is a network of operators in which the *maximum total response is attained by minimizing the maximum response over all of the individual operators.*

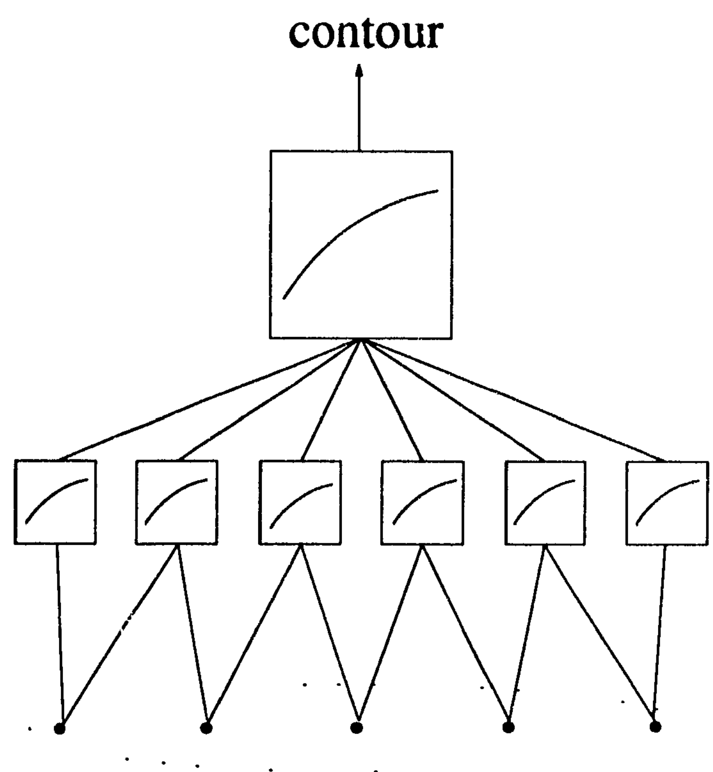


Figure 5.5: Nonlinear integration may lead to Minimax Gap Principle.

Chapter 6 Conclusion

Computer vision researchers have approached the problem of object perception from two directions. The inference of oriented structure from images, first assumed to be a trivial problem, turned out in fact to be quite hard, yet excellent progress has been made (Iverson & Zucker [1990]; Parent & Zucker [1989]). At the other extreme, while we are far from having a complete theory of shape, the last twenty years have witnessed some interesting ideas on shape languages and shape metrics (Biederman [1988]; Blum [1973]; Hoffman & Richards [1985]; Kimia, Tannenbaum & Zucker [1990]; Leyton [1989]; Mumford [1991]).

This leaves a bit of a gap. While there have been theories on the integration of tangent structure into more global contour representations (Blake & Zisserman [1987]; David & Zucker [1990]), I know of no computational theory for how such one-dimensional representations could lead to the computation of two- and three-dimensional shape. Probably the closest we have to a computational theory of contour closure is Ullman's sequential marking and tracing routine (Ullman [1984]), which involves tracing a continuous contour in search of either a termination point or a marked starting point.

While Ullman acknowledges that "fragmented contours can indeed often replace continuous ones" he does not explain how his visual routines could be extended to realistic image curves. The theory thus remains in the idealized domain of continuous mathematics. All of the interesting properties of perceptual closure: its continuous nature, its sensitivity to fragmentation geometry and contrast sign, its independence from region and surface completion processes, are not considered.

The deeper issue, however, is in the role that Ullman implicitly assumes for closure. The output of a closure routine based upon marking and tracing is a yes or a no. This is consistent with the feature view of closure which suggests that we simply want to register its presence or absence. But if we want to recognize an object or pick it up, we are not a great deal further along than when we started. In effect we have suspended shape processing until our closure routine terminates.

The computation of closure seems to involve geometric relationships at a range of scales. But what are these if not the elements of shape? The computation of shape representations and perceptual closure are intertwined. Just as for feature theory, the visual routine view of closure fails to make this critical link. The gap between the inference of oriented structure and the representation of shape has not been narrowed.

This gap in computational theory is paralleled by a gap in our knowledge of the physiology and anatomy of vision. While there is considerable debate about the best language in which to describe the function of striate cortex (Kulikowski & Kranda [1986]), there is no question that therein lies a highly organized representation of local, oriented image structure (Hubel & Wiesel [1968]), and that this representation is subject to modulation by the geometric context of the local image information (Nelson & Frost [1985]; Ts'o, Gilbert & Wiesel [1986]; Wiesel & Gilbert [1989]; Wiesel & Gilbert [1989]).

At the other extreme, there is considerable evidence for selectivity for faces, hands, and complex planar shapes in higher visual areas in temporal cortex (Desimone [1991]; Sakai & Miyashita [1991]; Tanaka et al. [1991]).

Between these extremes, our knowledge remains coarse. We know that there is a general progression in receptive field size, selectivity and complexity. There is ample evidence of attentional and behavioural modulations in areas V4 and IT (Maunsell et al. [1989]; Moran & Desimone [1985]; Sakai & Miyashita [1991]), and some claims for the generalization of oriented representations to include modally-completed contour (von der Heydt & Peterhans [1989]; Peterhans & von der Heydt [1989]). However, the biological substrate for the computation of shape from contour is at least as unknown as the computation itself.

I chose to study perceptual closure because I believe it lies near the core of this knowledge gap, both at the computational and the biological levels, and because I believe that the role of closure in perception has been misperceived both by psychologists and by computational theorists.

If we are to understand how we perceive shape and recognize objects, we must stop reducing ideas with high computational potential to simple yes or no questions. Rather than asking "Is closure a feature" we must ask "How does perceptual closure advance our computations of shape?"

This work begins with the assumption that a visual system represents object shape, and the observation that shape can be wonderfully depicted by contour. While computational models for this perceptual process are still rather thin, what we have relies heavily on two-dimensional properties such as curvature sign. Since we know that mathematically the computation of these properties depends on the contour being closed, it seems vital to study the meaning of contour closure in a perceptual context, with a view toward the representation of multi-dimensional shape.

My experiments go beyond supporting the basic hypothesis of a perceptual closure continuum mediating shape perception, to revealing several properties of perceptual closure and principles of shape from contour. A Minimax Gap Principle constrains the dependence of perceptual closure on contour fragment geometry. A Contrast Sign Principle constrains the photometric sensitivity of perceptual closure. My experiments elucidate the relationship between perceptual closure and properties of connectedness, enclosure and symmetry, the distinction between boundary and region processes, and the separate existence of early grouping process and later surface completion processes leading to the perception of multiple occlusion.

The exploration of a functional basis for my results leads to an interpretation of perceptual closure as a measure of the confidence with which contours can be interpreted as projections from a single object boundary. This leads to an equivalence relation between perceptual closure and *single object confidence.*

Neurons in temporal cortex of primate selective for shape properties such as concavity and convexity have been reported (Tanaka et al. [1991]). Computationally, these properties depend upon contour closure, and I have characterized a psychophysical manifestation of this dependence. It would be very interesting to now seek a physiological basis for perceptual closure.

Appendix A An Unbiased Visual Search Methodology

In the traditional visual search technique (Treisman & Gelade [1980]), subjects are shown an equal mix of displays with a target and displays without a target. Subjects press one button when they have detected a target in the display, another if they are unable to find a target. But if a subject is having problems finding a target, how does he or she decide when to stop looking? One suggestion is that subjects use a rough timing mechanism based upon how difficult they expect the task to be (Chun, Wolfe & Friedman-Hill [1991]).

Although subjects rarely indicate that a target is present when it is not (target absent error rates are seldom more than 10% and typically average about 5%), it is much more common for subjects to indicate that the target is absent when in fact it is present in the display (target present error rates often average more than 10% and can be as high as 20% or even 30%) (Donnelly, Humphreys & Riddoch [1991]; Enns & Rensink [1991]). Moreover, this error rate has been observed to correlate positively with reaction time (Enns & Rensink [1991]). Thus, to use the timer analogy, when the task is difficult, subjects time-out more often.

Of course when an error is made, the trial is not used in determining the mean reaction time for the associated stimulus and display size. *This means that these estimates will be systematically depressed.* Worse yet, this depression will be greater for conditions which are harder: larger display sizes with less easily discriminable stimuli. This results in a general depression of search slope estimates, particularly for more difficult tasks.

The size of the variance under particular stimulus conditions is normally quite large in visual search experiments. For example, second-order fits to the response-time data of my experiments indicate that less than one third of the variance for a particular stimulus type is due to variation in display size. Moreover, this unmodeled variance increases with response time. This means that the biases caused by the disqualification of hard trials could be very significant.

The visual search technique used in this work overcomes the limitations of the classical procedure. This is accomplished primarily by eliminating the uncertainty associated with classical visual search: in my technique the subjects know that the target will always be *somewhere* in the display. What they don't know is *where*. Thus subjects keep looking until they find the target: they never time-out.

By removing this dominant source of error in target present displays, I have achieved error rates averaging less than 2%. Thus even if there remains some small correlation between error rate and response time, the error rate is too low for the resulting bias to be significant.

In order to both validate the technique and test for bias, I repeated the basic closure experiment of section 4.3 using the classical procedure. The results (Fig. A.1) confirm that search is rapid for closed stimuli but slow for open stimuli.[1]

Table A.1 shows the linear model parameters for the results using the classical procedure and the results using my procedure (section 4.3). Fig. A.2 shows the target-present results for the classical procedure plotted with the results using my procedure. While the results for the closed

[1]Search slope for the closed stimuli is significantly less than that for the open stimuli ($p < 0.005$). Intercepts do not differ significantly ($p > 0.1$)

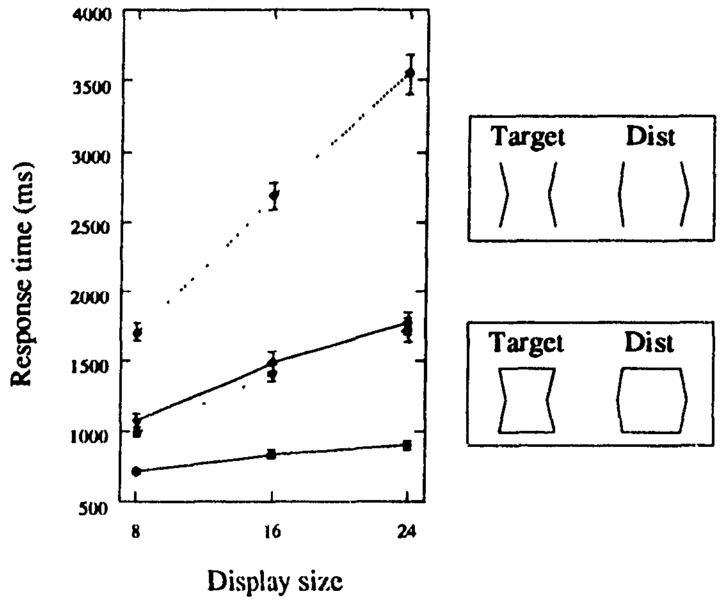


Figure A.1: Visual search results using the classical procedure (14 subjects). Target-present results are shown solid, target-absent results are shown dotted.

stimuli do not differ significantly, search for the open figures appears to be faster using the classical procedure.[2]

An examination of the error rates for these experiments may shed light on this difference. The error rates using our technique remained under 3%, averaging 1% for the closed stimuli and 1.8% for the open stimuli. The error rates using the classical technique are shown in Fig. A.3. Again, error is below 3% *except for the target-present condition of the difficult trials*: display sizes of 16 and 24 with open stimuli resulted in error rates of 7.7% and 10.2% respectively.

Thus, as predicted, the classical technique leads to target-present

[2]For the closed figures, neither slope nor intercept differ significantly between methods ($p > 0.1$) For the open figures, both slope ($p < 0.005$) and intercept ($p < 0.05$) are significantly lower for the classical procedure than for my procedure.

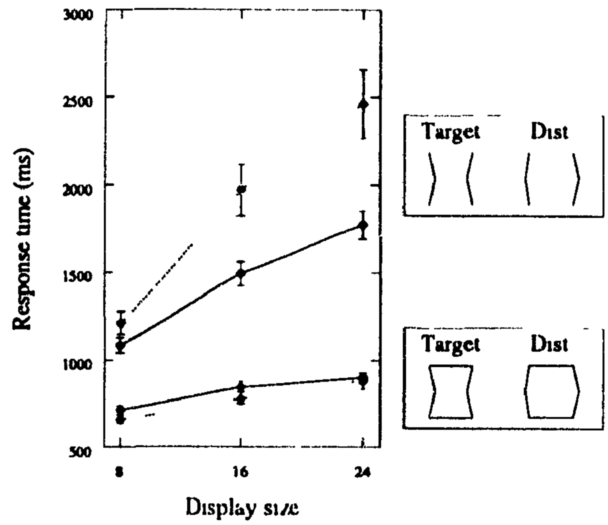


Figure A.2: Comparison of visual search results using different experimental procedures. Results using classical procedure are shown solid. Results using my procedure are shown dotted.

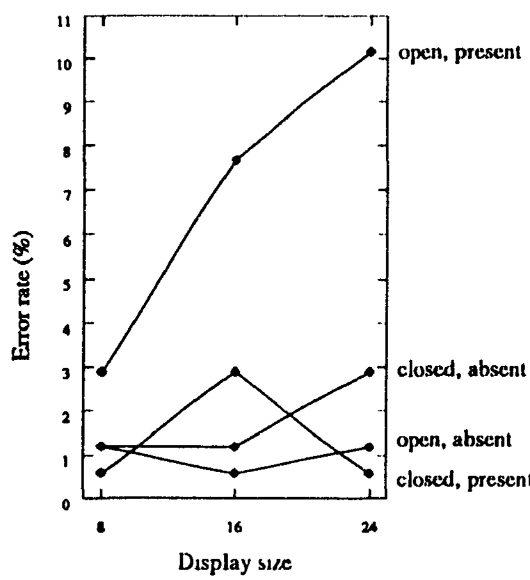


Figure A.3: Mean error rates for the classical visual search procedure

Procedure	Closure	Target Status	Slope (ms/item)	Intercept (ms)
Classical	Closed	Present	12	621
Classical	Closed	Absent	48	609
Classical	Open	Present	45	727
Classical	Open	Absent	116	786
Unbiased	Closed	Present	14	546
Unbiased	Open	Present	83	555

Table A.1: Linear fit parameters for search results using classical procedure

error rates which are strongly correlated with response time and which become significant for the large display sizes of hard tasks. I believe that this accounts for the depressed search slope observed for the open stimuli using the classical visual search procedure (Fig. A.2).

References

F. Attneave [1954], "Some Informational Aspects of Visual Perception.," *Psychol. Rev.* 61, 183-193.

I. Biederman [1988], "Aspects and Extensions of a Theory of Human Image Understanding," in *Computational Processes in Human Vision*, Z. Pylyshyn, ed., Ablex, Norwood, NJ, 370 427.

A. Blake & A. Zisserman [1987], *Visual Reconstruction*, MIT Press, Cambridge, Mass..

H. Blum [1973], "Biological Shape and Visual Science (Part I)," *J. Theor. Biol.* 38, 205–287.

P.J. Burt & E.H. Adelson [1983], "The Laplacian Pyramid as a Compact Image code," *IEEE Trans. Commun.* COM-31, 532 540.

T. Caelli & B. Julesz [1978], "On Perceptual Analyzers Underlying Visual Texture Discrimination: Part I," *Biol. Cybernetics* 28, 167 175.

T.M. Caelli, B. Julesz & E.N. Gilbert [1978], "On perceptual analyzers underlying visual texture discrimination: Part II," *Biol. Cybern.* 29, 201–214.

M.P. do Carmo [1976], *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, NJ.

P. Cavanagh & Y.G. Leclerc [1989], "Shape from Shadows," *J. Exp. Psychol.: H.P.&P* 15, 3-27.

M. M. Chun, J. M. Wolfe & S. R. Friedman-Hill [1991], "Just Say No: How to Terminate a Visual Search," *J. Invest. Opthalm. Visual Sci.* 32, 1041.

K. Clark [1956], *The Nude: A Study in Ideal Form*, Princeton University Press, Princeton, NJ.

C. David & S.W. Zucker [1990], "Potentials, valleys and dynamic global coverings," *Int. J. Computer Vision* 5, 219–238.

R. Desimone [1991], "Face-Selective Cells in the Temporal Cortex of Monkeys," *J. Cognitive Neuroscience* 3, 1–8.

N. Donnelly, G. W. Humphreys & M. J. Riddoch [1991], "The Parallel Computation of Primitive Shape Descriptions," *J. Exp. Psychol.: H.P.&P* 17, 561–570.

J.T. Enns & R. A. Rensink [1991], "Preattentive Recovery of Three-Dimensional Orientation from Line Drawings," *Psychol. Rev.* 98, 335–351.

L. Glass & E. Switkes [1976], "Pattern recognition in humans: correlations which cannot be perceived," *Perception* 5, 67–72.

S. Grossberg & E. Mingolla [1985], "Neural Dynamics of Form Perception: Boundary Completion, Illusory Figures, and Neon Color Spreading," *Psychol. Rev.* 92, 173–211.

R. Gurnsey & R.A. Browse [1989], "Asymmetries in visual texture discrimination," *Spatial Vision* 4, 31–44.

P. Hammond & D.M. Mackay [1983], "Influence of luminance gradient reversal on simple cells in feline striate cortex," *J. Physiol.* 337, 69–87.

P. Hammond & D.M. Mackay [1985], "Influence of luminance gradient reversal on complex cells in feline striate cortex," *J. Physiol.* 359, 315–329.

M. Henle [1979], *A Combinatorial Introduction to Topology*, W.H. Freeman, San Francisco.

R. von der Heydt & E. Peterhans [1989], "Mechanisms of Contour Perception in Monkey Visual Cortex. I. Lines of Pattern Discontinuity ," *J. Neuroscience* 9, 1731–1748.

D. D. Hoffman & W. A. Richards [1985], "Parts of Recognition," *Cognition* 18, 65–96.

D.H. Hubel & T.N. Wiesel [1960a], "Receptive fields of optic nerve fibres in the spider monkey," *J. Physiol.* 154, 572–580.

D.H. Hubel & T.N. Wiesel [1960b], "Single unit activity in lateral geniculate body and optic tract of unrestrained cats," *J. Physiol.* 150, 91-104.

D.H. Hubel & T.N. Wiesel [1962], "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.

D.H. Hubel & T.N. Wiesel [1968], "Receptive fields and functional architecture of monkey striate cortex.

L.A. Iverson & S.W. Zucker [1990], "Logical/Linear Operators for Measuring Orientation and Curvature," McGill Univ., Tech. Rep. TR-CIM-90-06, Montréal, Canada.

B. Julesz [1962], "Visual Pattern Discrimination," *IRE Trans. on Information Theory* IT-8, 84–92.

B. Julesz [1980], "Spatial nonlinearities in the instantaneous perception of textures with identical power spectra," *Phil. Trans. Roy. Soc. Lond., Ser. B* 290, 83-94.

B. Julesz [1981], "Textons, the elements of texture perception and their interactions," *Nature* 290, 91-97.

B. Julesz [1986], "Texton Gradients: The Texton Theory Revisited," *Biol. Cybern.* 54, 245–251.

B. Julesz [1991], "Early vision and focal attention," *Rev. Modern Phys.* 63, 735–772.

B. Julesz & J.R. Bergen [1987], "Textons: The Fundamental Elements in Preattentive Vision and Perception of Textures," in *Readings in Computer Vision*, M.A. Fischler & O. Firschein, eds., Morgan Kaufmann, Los Altos, CA, 243–256.

B. Julesz, E.N. Gilbert, L.A. Shepp & H.L. Frisch [1973], "Inability of humans to discriminate between visual textures that agree in second-order statistics - revisited," *Perception* 2, 391–405.

G. Kanizsa [1979], *Organization in Vision*, Praeger, New York.

B. Kimia, A. Tannenbaum & S. W. Zucker [1990], "Toward a computational theory of shape: An overview," *Proc. 1st European Conf. on Computer Vision* 427, 402–407.

J.J. Koenderink [1984], "What does the occluding contour tell us about solid shape?," *Perception* 13, 321–330.

K. Koffka [1935], *Principles of Gestalt Psychology*, Harcourt, Brace & World, New York.

S.W. Kuffler [1953], "Discharge patterns and functional organization of mammalian retina," *J. Neurophysiol.* 16, 37–68.

J.J. Kulikowski & K. Kranda [1986], "Image Analysis performed by the visual system: feature *versus* Fourier analysis and adaptable filtering," in *Visual neuroscience*, J.D. Pettigrew, ed., Cambridge, UK, 381–404.

J.Y. Lettvin, H.R. Maturana, W.S. McCulloch & W.H. Pitts [1959], "What the frog's eye tells the frog's brain," *Proc. Inst. Radio Engrs.* 47, 1940–1951.

M. Leyton [1989], "Inferring Causal History from Shape," *Cognitive Science* 13, 357–387.

D.G. Lowe [1985], *Perceptual Organization and Visual Recognition*, Kluwer, Boston.

J.M.A. Malik & P. Perona [1990], "Preattentive texture discrimination with early vision mechanisms," *J. Opt. Soc. Am.* 7, 923–932.

D. Marr [1982], *Vision*, W.H. Freeman, New York.

J.H.R. Maunsell, T.A. Nealey, G. Sclar & D.D. DePriest [1989], "Representation of Extraretinal Information in Monkey Visual Cortex," in *Neural Mechanisms of Visual Perception*, D. Lam & C. Gilbert, eds., Gulf Publishing Company, Houston, Texas, 223–235.

J. Moran & R. Desimone [1985], "Selective Attention Gates Visual Processing in the Extrastriate Cortex," *Science* 229, 782–784.

D. Mumford [1991], "Mathematical Theories of Shape: Do they model perception," *Geometric Methods in Computer Vision (SPIE)* 1570.

U. Neisser [1969], *Cognitive Psychology*, Appleton-Century-Crofts, New York.

J.I. Nelson & B.J. Frost [1985], "Intracortical facilitation among co-oriented, co-axially aligned simple cells in cat striate cortex," *Exp. Brain. Res.* 61, 54–61.

K. Nicolaïdes [1941], *The Natural Way to Draw*, Houghton Mifflin Company, Boston.

M. Nitzberg [1991], "Depth from Overlap," Harvard Univ., PhD Thesis.

P. Parent & S.W. Zucker [1989], "Trace Inference, Curvature Consistency, and Curve Detection," *IEEE Trans. Pattern Anal. Machine Intell.* 11, 823–839.

P. Perona & J. Malik [1990], "Detecting and localizing edges composed of steps, peaks and roofs," *Proc. 3rd Int. Conf. on Computer Vision*, Osaka, Japan.

E. Peterhans & R. von der Heydt [1989], "Mechanisms of Contour Perception in Monkey Visual Cortex. II. Contours Bridging Gaps," *J. Neuroscience* 9, 1749–1763.

J.R. Pomerantz, L.C. Sager & R.G. Stoever [1977], "Perception of wholes and their component parts: Some configural superiority effects," *J. Exp. Psychol.: H.P.&P.* 3, 422–435.

K. Prazdny [1983], "Illusory contours are not caused by simultaneous brightness contrast," *Perception & Psychophysics* 34, 403–404.

K. Prazdny [1986], "Psychophysical and computational studies of random-dot Moire patterns," *Spatial Vision* 1, 231–242.

I. Rock [1983], *The Logic of Perception*, MIT Press, Cambridge, Mass..

I. Rock & S. Palmer [1990], "The Legacy of Gestalt Psychology," *Scientific American* 263, 84–91.

K. Sakai & Y. Miyashita [1991], "Neural organization of the long-term memory of paired associates," *Nature* 354, 152–155.

R. Shapley & J. Gordon [1985], "Nonlinearity in the perception of form," *Perception & Psychophysics* 37, 84–88.

K. Tanaka, H. Saito, Y. Fukada & M. Moriya [1991], "Coding Visual Images of Objects in the Inferotemporal Cortex of the Macaque Monkey," *J. Neurophysiol.* 66, 170–189.

A. Treisman & G. Gelade [1980], "A feature integration theory of attention," *Cog. Psychol.* 12, 97–136.

A. Treisman & S. Gormican [1988], "Feature analysis in early vision: Evidence from search asymmetries," *Psychol. Rev.* 95, 15–48.

A. Treisman & R. Paterson [1984], "Emergent Features, Attention, and Object Perception," *J. Exp. Psychol.: H.P.&P.* 10, 12–31.

A. Treisman & H. Schmidt [1982], "Illusory Conjunctions in the Perception of Objects," *Cog. Psychol.* 14, 107–141.

A. Treisman & J. Souther [1985], "Search Asymmetry: A Diagnostic for Preattentive Processing of Separable Features," *J. Exp. Psychol.: Gen.* 114, 285–310.

D.Y. Ts'o, C.D. Gilbert & T.N. Wiesel [1986], "Relationships Between Horizontal Interactions and Functional Architecture in Cat Striate Cortex as Revealed by Cross-Correlation Analysis," *J. Neuroscience* 6, 377–385.

S. Ullman [1984], "Visual routines," *Cognition* 18, 298–328.

T.N. Wiesel & C.D. Gilbert [1989], "Neural Mechanisms of Visual Perception," in *Neural Mechanisms of Visual Perception*, D. Lam & C. Gilbert, eds., Gulf Publishing Company, Houston, Texas, 7–33.

A. Witkin [1983], "Scale Space Filtering," *Proc. Int. Joint Conf. on Artif Intell.*, Karlsruhe.

A. Witkin, D. Terzopoulos & M. Kass [1987], "Signal Matching Through Scale Space," *Int. J. Computer Vision* 1, 133–144.

A.P. Witkin & J.M. Tenenbaum [1983], "On the role of structure in vision," in *Human and machine vision*, J. Beck, B. Hope & A. Rosenfeld, eds., Academic Press, New York, 481–583.

S.W. Zucker [1986], "The Diversity of Perceptual Grouping," in *Vision, Brain and Cooperative Computation*, M. Arbib & A. Hanson, eds., MIT, Cambridge, Mass., 231-261.

S.W. Zucker & S. Davis [1988], "Points and Endpoints: A size/spacing constraint for dot grouping," *Perception* 17, 229-247.

S.W. Zucker, K.A. Stevens & P. Sander [1983], "The relation between proximity and brightness similarity in dot patterns," *Perception & Psychophysics* 34, 513–522.