

## The topological approach to perceptual organization

Lin Chen

*Key Laboratory of Cognitive Science, Graduate School and Institute of Biophysics, Chinese Academy of Sciences, Beijing, China*

To address the fundamental question of “what are the primitives of visual perception”, a theory of topological structure and functional hierarchy in visual perception has been proposed. This holds that the global nature of perceptual organization can be described in terms of topological invariants, global topological perception is prior to the perception of other featural properties, and the primitives of visual form perception are geometric invariants at different levels of structural stability. In Part I of this paper, I will illustrate why and how the topological approach to perceptual organization has been advanced. In Part II, I will provide empirical evidence supporting the *early* topological perception, while answering some commonly considered counteraccounts. In Part III, to complete the theory, I will apply the mathematics of tolerance spaces to describe global properties in discrete sets. In Part IV, I will further present experimental data to demonstrate the *global-to-local* functional hierarchy in form perception, which is stratified with respect to structural stability defined by Klein’s Erlangen Program. Finally, in Part V, I will discuss relations of the *global-to-local* topological model to other theories: The topological approach reformulates both classical Gestalt holism and Gibson’s direct perception of invariance, while providing a challenge to computational approaches to vision based on the *local-to-global* assumption.

### INTRODUCTION

#### A great divide: Local-to-global vs. global-to-local

As a Chinese proverb says: Everything is difficult at its very beginning. Historically, major schools of vision diverge in their answers to the question of “Where visual processing begins?” (Pomerantz, 1981) or “What are the primitives of visual perception?” (Chen, 1982). The question is so fundamental and also so controversial as to serve as a watershed, a Great Divide, separating

---

Please address all correspondence to: Lin Chen, Key Laboratory of Cognitive Science, Graduate School and Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Beijing 100101, China. Email: chenl@intr.nimh.nih.gov

Supported by National Nature Science Foundation of China Grant (697900800); Ministry of Science and Technology of China Grant (1998030503); Chinese Academy of Sciences Grants (KGCX2-SW-101, KJCX1-07). This work was partly done during author’s sabbaticals, at Institute of Medical Psychology, University of Munich, and at National Institute of Mental Health.

two most basic and sharply contrasting lines of thinking in the study of perception.

*Early feature analysis: Local-to-global.* On one side of the Great Divide, the early feature-analytic viewpoint holds that perceptual processing is *from local to global*: Objects are initially decomposed into separable properties and components, and only in subsequent processes are objects recognized, on the basis of the extracted features. The computational approach to vision by Marr (1982), representative of “early feature-analysis” viewpoint, claims that the primitives of visual-information representation are simple components of forms and their local geometric properties, such as, typically, line segments with slopes.

*Early holistic registration: Global-to-local.* On the other side of the Great Divide, the viewpoint of early holistic registration claims that perceptual processing is *from global to local*: Wholes are coded prior to perceptual analysis of their separable properties or parts, as indicated by the conception of perceptual organization in Gestalt psychology. As we will see in the following discussion, with respect to the fundamental question of “Where to begin”, the core contribution of the Gestalt idea goes far beyond the notion that “Whole is more than the simple sum of its parts”; rather it is that “Holistic registration is prior to local analysis”.

The idea of early feature analysis has gained wide acceptance, and dominates much of the current study of visual cognition. Intuitively, it seems to be natural and reasonable that visual processing begins with analysing simple components and their local geometric properties, typically as line segments with slopes, as they are readily to be considered physically simple and computationally easy. An underlying idea of Marr’s computational system of vision was, in Marr’s own words, “In the early stages of the analysis of an image, the representations used depend more on what it is possible to compute from an image than on what is ultimately desirable.” (Marr, 1978). Nevertheless, a starting point of the present paper is that physically or computationally simple doesn’t necessarily mean psychologically simple or perceptually primitive; therefore, the question of which variables are perceptual primitives is not a question answered primarily by logical reasoning or analysis of computational complexity but rather by empirical findings.

### Topological structure and functional hierarchy in form perception

To address the fundamental question of what are the primitives of visual perception, based on a fairly large set of data on perceptual organization reviewed here, a theory of “*early topological perception*” has been proposed. This holds that:

A primitive and general function of the visual system is the perception of topological properties. The time dependence of perceiving form properties is systematically related to their structural stability under change, in a manner similar to Klein's hierarchy of geometries; in particular, topological perception (based on physical connectivity) is prior to the perception of other geometrical properties. The invariants at different geometrical levels are the primitives of visual form perception. These include, in a descending order of stability (from global to local), topological, projective, affine, and Euclidean geometrical invariants.

The topological approach is based on one core idea and includes two main aspects. The core idea is that perceptual organization should be understood in the perspective of transformation and perception of invariance over transformation. The first aspect emphasizes the topological structure in form perception, namely, that the global nature of perceptual organization can be described in terms of topological invariants. The second aspect further highlights *early* topological perception, namely, that topological perception is prior to perception of local featural properties. The "prior" has two strict meanings: First, it implies that global organizations, determined by topology, are the basis that perception of local geometrical properties depends on; and second, topological perception (based on physical connectivity) occurs *earlier* than the perception of local geometrical properties.

The hypothesis of *early* topological perception led to a major finding that the relative perceptual salience of different geometric properties is remarkably consistent with the hierarchy of geometries according to **Klein's Erlangen Program** (see Part II and IV), which stratifies geometries in terms of their relative stability over transformations. Based on the finding, a functional hierarchy in form perception has been established as a formal and systematic definition of "*global-to-local*" relations: A property is considered more global (or stable) the more general the transformation group is, under which this property remains invariant; relative to geometrical transformations, the topological transformation is the most general and hence topological properties are the most global.

The framework of the topological structure and functional hierarchy highlights a fundamental empirical prediction, namely a time dependence of perceiving form properties, in which visual processing is *from global to local*: The more global a form invariant is the earlier its perception occurs, with topological perception being the most global and occurring earliest. The framework further highlights a series of novel empirical predictions about long-standing issues related to the study of perceptual organization, and many Gestalt-type phenomena in form perception may be explained in this unified manner. They include the following examples:

- With respect to the relationship between different organizational factors, proximity is the most fundamental organizational factor (even in comparison with uniform connectivity; Palmer & Rock, 1994) (see Part III), and there is a time course of processing different organizational principles: Proximity precedes similarity, and topological similarity precedes similarity of local geometric properties (see Part VI).
- Early topological perception predicts the visual sensitivity to distinction made in topology. For example, **two stimuli that are topologically different are more discriminable under a near-threshold condition than are other pairs of forms that are topologically equivalent despite their difference in local features** (see Part II).
- With respect to the question of whether motion perception precedes form perception or vice versa, topological discrimination should occur earlier than and determine motion perception (see Part II).
- Configural superiority effects (Pomerantz, 1981) demonstrated by configural relations between line segments, such as the “triangle–arrow pair”, may simply demonstrate the superiority effect for perception of holes over individual line segments (see Part IV).
- With respect to “global precedence” (Navon, 1977), according to the functional hierarchy, the performance advantage for compound letters is quite straightforward: Global precedence reflects the primacy of proximity in perceptual organization (see Part III).
- If topological properties are primitive, illusory conjunctions (Treisman & Gelade, 1980) of topological properties, such as holes, should sometimes take place (see Part II).
- With respect to the definition of perceptual object, the topological approach **ties a formal definition of “object” to invariance over topological transformation** (see Part I). From this definition, it follows that as an object is moving along and a hole appears in it, this topological change would disturb object continuity, while changes of shape and colour wouldn’t (Wolfe, personal communication). For example, in an MOT (multiple object tracking) test (Pylyshyn & Storm, 1988; vanMarle & Scholl, 2003), attentive tracking processes would be impaired by objects changing topology by getting a hole, while it does not matter if they change local features and colours.
- With respect to its ecological function and functional anatomy, long-range apparent motion works by abstracting form invariants, and hence is associating with form perception and activates the ventral pathway in the two visual systems (Ungerleider & Mishkin, 1982). Specifically, the fMRI activation should be correlated with the form stability under change (see Parts II and IV).
- From the perspective of biological evolution, if topological perception is indeed a fundamental property of vision, one might expect topological



properties to be extracted by all visual systems, including the relatively simple ones possessed by insects, such as bees (see Part II).

In summary, the framework of topological structure and functional hierarchy in form perception provides a new analysis of the fundamental question, i.e., “What are the primitives of visual perception?”, in which primitives of visual form perception are considered to be geometric invariants (as opposed to simple components of objects, such as line segments) at different levels of structural stability.

In the following, I will illustrate why and how the topological approach to perceptual organization has been advanced (Part I); provide empirical evidence supporting the topological perception, while answering some commonly considered counteraccounts (Part II); complete the theory of topological perception, using the mathematics of tolerance spaces that describe global properties in discrete sets (Part III); present experimental data to demonstrate the functional hierarchy in form perception, which is stratified with respect to structural stability defined by Klein’s Erlangen Program (Part IV); and finally, discuss relations of *early* topological perception to other theories, including Gestalt psychology, Gibsonian psychology, and the computational approach (Part V).

## PART I: WHY AND HOW—A TOPOLOGICAL APPROACH TO PERCEPTUAL ORGANIZATION

### A paradoxical problem of “where to put the master map”

Fundamental problems faced by the early feature-analysis approach are typically embodied in a paradoxical problem of “where to put the master map” as posed by the feature-integration theory of Treisman and co-workers (e.g., Treisman & Gelade, 1980). Feature-integration theory, consistent with the early feature-analysis approach, adopts a “two-stage model”: In the first, preattentive stage, primitive features, such as colours and orientations, are abstracted effortlessly and in parallel over the entire visual field, and registered in special modules of feature maps; and in the second, attentive stage, focal attention is required to recombine the separate features to form objects. A master map of locations plays a central role in feature binding by tying the separate feature maps together. Within the master map, a focal attention mechanism selects a filled location, binding the activated features linked to that location together to form a coherent object.

Problems for feature-integration theory are, however, represented by the question of where exactly the master map of locations fits into the feature integration mechanism. In Treisman’s own words, “I have hedged my bets on where to put the master map of locations by publishing two versions of the figure! In one of them, the location map received the output of the feature

modules (e.g., Treisman, 1986) and in other is placed at an earlier stage of analysis (e.g., Treisman, 1985; Treisman & Gormican, 1988)' (Treisman, 1988, pp. 203–204).

To place the master map of locations at an early stage of analysis, in Treisman's own words, "implies that different dimensions are initially conjoined in a single representation before being separately analysed" (Treisman, 1988, pp. 203–204). This contradicts the basic position of early feature analysis. Placing the master map later, however, contradicts some behavioural data. One of the strengths of feature-integration theory is that it draws on a number of major pieces of counterintuitive evidence, including illusory conjunction and visual search, which appear to provide strong support for early feature analysis (e.g., Treisman & Gelade, 1980; Treisman & Gormican, 1988). However, it is interesting to see that problems for this theory also arise here (e.g., for a general review, see Humphreys & Bruce, 1989). For example, despite the fact that line segments are commonly considered to be basic features, there is markedly little evidence for illusory conjunction where line segments are miscombined into letter-like objects, when letter shapes and line segments forming the letter shapes are used as stimulus forms (e.g., Duncan, 1984). In contrast, there is much stronger evidence that whole letter shapes migrate across words and produce illusory conjunctions of the entire letter shapes, rather than of line segments making up the letter shapes. These findings indicate that letter shapes, as combinations of line segments, behave psychologically as holistic objects, even though line segments are commonly considered to be basic features. Apparently attention, as it relates to feature binding, is not needed for holistic object perception. This suggests that before a stage of separate featural analysis, there must be a stage of early holistic perception in which objects like letters are coded as wholes.

Treisman and co-workers, in their effort to explain these counterexamples, have augmented feature-integration theory with new strategies and new mechanisms of attention, such as "guided search" (e.g., Wolfe, 1994), "map suppression", and dividing items into different depth planes. Nevertheless, these efforts are not completely successful (e.g., Duncan & Humphreys, 1989) but rather in fact illustrate that, despite the attractions of feature-integration theory, the paradoxical problem of "Where to put the master map" stems directly from the fundamental question of "Where visual processing begins".

### **Perceptual organization: To reverse back the inverted (upside-down) problem of feature binding**

Regardless of how an object is decomposed into properties and components, the decomposed features themselves are unlikely to be sufficient for achieving object recognition. Indeed, we do not normally perceive isolated features such as brightness, colours, and orientations free from an object, leading to the contention that there must be a further process of feature binding. This problem of

feature binding presents a central problem for current vision research in particular, and for parallel and distributed modelling of cognition in general (e.g., Müller, Elliott, Herman, & Mecklinger, 2001).

However, despite the centrality of the issue for perceptual theory, it is questionable whether any breakthrough has been made after extensive efforts. Both spatial and temporal factors have been considered as cues for binding features together. But the principles for feature binding based on either space or time are neither always obeyed nor exclusive.

Feature binding and perceptual organization appear to be very similar problems (Duncan, 1989) in the sense that both deal with similar processes, such as “what goes with what”, and with similar concepts, such as belongingness and assignment. It turns out that, even though the early feature-analysis viewpoint emphasizes the fundamental importance of early parallel processing, the issue of perceptual organization remains indispensable. Yet, the concepts of “perceptual organization” and “feature binding” involve very different underlying issues, with the former rooted in the idea of early holistic registration and the latter originating from an assumption of early feature analysis. Thus, with respect to the fundamental question of “Where to begin”, perceptual organization and feature binding can be considered contrary concepts, going in opposite directions.

In terms of our understanding of objects in the real world, there may be little disagreement that the real features of an object, whether geometrical or physical, exist together as a coherent whole belonging to one entity in the outside world. The question of how the perceptual system represents perceptual objects as fundamental units of conscious perceptual experience, however, has either given rise to much controversy when considered, or not been considered at all. But the truth remains that real features of a real object, at a given time, originally coexist together rather than being separated; a real object is an integral stimulus, a single thing. This truth is a fundamental property of a real-world object. No one doubts the direct perception of various featural properties such as brightness, colour, line orientation, and so on. Why, then, is only this fundamental property of “belonging together as a whole” excluded from the membership of primitives in our perceptual world? The assumption that the visual system cannot directly perceive a real integral object has not yet been proved or disproved. Indeed, the continuing challenges to the issues of feature binding suggest that this question deserves closer scrutiny.

From the perspective of early holistic registration, the feature-binding problem is an ill-posed question: Not just a question of getting off on a wrong foot but even a question of “standing upside down”. In this sense, the feature-binding problem might be a wrong, inverted question. Kubovy and Pomerantz (1981) commented: “the main problems facing us today are quite similar to those faced by the Gestalt psychologists in the first half of this century”. After half a century, the study of visual perception appears, in some sense, to be back

to square one. This situation leads us to wonder whether the problems of feature binding are due to difficulties in posing the fundamental question of “Where to begin”.

Where does the above analysis leave us? It suggests that early holistic registration may provide a way to avoid the feature-binding problem by focusing on issues of perceptual organization. In other words, we may apply the concept of perceptual organization to reverse back the inverted (upside-down) question of feature binding.

### The topological approach

Despite its deep and rational core in the idea of early holistic registration, the notion of perceptual organization has its own problems. In particular, like other Gestalt concepts, it has suffered from a lack of proper theoretical treatment. Gestalt evidence has often been criticized for being mainly phenomenological and relying mainly on conscious experience. Consequently, explanations from theories of perceptual organization usually rely on intuitive or mentalistic concepts that are somewhat vague and elusive. What is needed is a proper formal analysis of perceptual organization that goes beyond intuitive approaches, and provides a theoretical basis for describing or defining precisely the core concepts related to perceptual organization, e.g., “global” vs. “local”, “objects”, “grouping”, and others. Until the intuitive notions of these Gestalt-inspired concepts become properly and precisely defined, the proposed principles of perceptual organization cannot be entirely testable.

#### *Delimiting the concept of perceptual organization*

To give a precise description of the essence of perceptual organization, we first need to properly delimit the concept of perceptual organization. On the one hand, as Rock (1986) pointed out, “The concept of perceptual organization should not be defined so loosely as to be a synonym of perception”; on the other hand, this concept should not be so limited as to be unable to cover the great variety and the commonplace occurrence of perceptual organization. The following definition of perceptual organization given by Rock (1986) is considered to define properly the very notion of perceptual organization:

The meaning of organization here is the grouping of parts or regions of the field with one another, the “what goes with what” problem, and the differentiation of figure from ground.

According to this definition, the study of perceptual organization is concerned with early stages of perceptual processing divorced from high-level cognition, and therefore such delimitation pitches our discussion at the right level to answer the basic question of “Where visual processing begins”. On the other hand, the

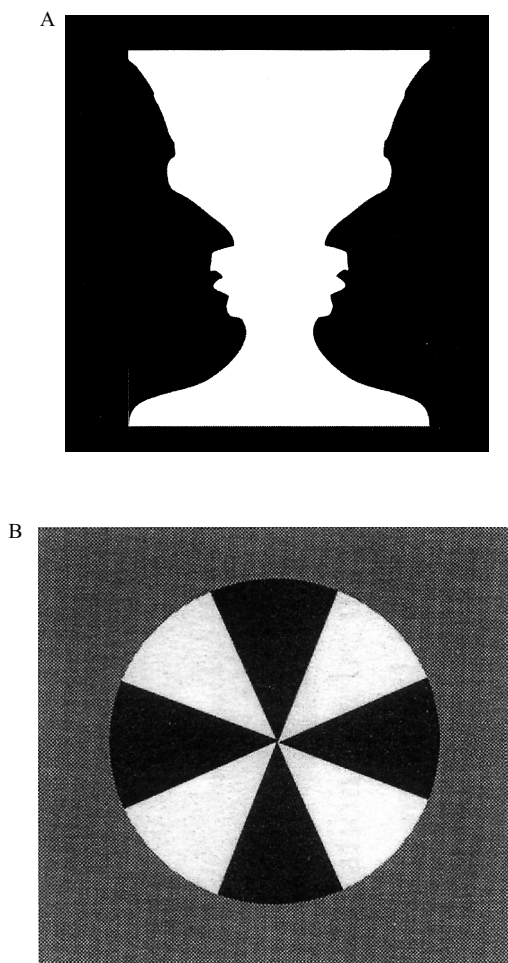


concept of perceptual organization discussed in the present paper deals with general processes, such as figure–ground differentiation, grouping, “what goes with what”, belonging and assignment (not particular processes, such as differentiating luminance flux, discriminating orientation, or recognizing a face), and with abstract things, such as objects, units, and wholes as well as their counterparts, such as items, elements, and parts (not concrete things, such as a line segments, a geometrical figure, a friend’s face). These general processes and abstract things represent the variety and commonplace occurrence of perceptual organization. Figure 1 illustrates this. The percept of Figure 1A may be described at a semantic level: Either a vase or two profiles face to face. On the other hand, it may be described in terms of the vocabulary of perceptual organization: Two boundaries (units) grouped into one object (as the basis of the percept of a vase) or two boundaries (units) separated into two objects (as the basis of the percept of the two profiles face to face). It is the latter level, the level of perceptual organization, which our present research focuses on. Furthermore, as Figure 1B demonstrates, perceptual organization may be perceived without semantic meaning. Here even though there is little semantic meaning involved in the stimulus, either the black parts are perceived to be an unified whole as a figure and the grey parts, another unified whole as background, or vice versa. Top-down processing of prior knowledge or expectation may influence perceptual organization, but it will avoid possible confusion if we consider perceptual organization and top-down processing of high-level cognition separately. This strengthens the rationale for defining the terminology for describing perceptual organization, emphasizing the primacy of perceptual organization.

*Major challenges to establishing a proper theoretical treatment on perceptual organization: Its commonplace, and its general and abstract characteristics*

The concept of perceptual organization reflects the most common fact that the phenomenal world contains objects separated from one another by space or background. Phenomena in perceptual organization are usually so common that they have not been looked on as an achievement of the perceptual system, and, thus, as something to be explained (Rock, 1986). For example, tremendous efforts have been made to study how to detect line segments with orientation and location, but little attention has been paid to the question of how to perceive a line segment as a single object. While the study of face recognition has advanced considerably, the fundamental grouping question of “which eyes go with which noses, which noses with which mouths, and so forth” (Pomerantz, 1981) has been almost completely ignored.

One more example shows how commonplace characteristics of perceptual organization make it difficult to realize that there are problems requiring



**Figure 1.** (A) An ambiguous figures of “a vase vs. two faces”, showing competing organization. (B) An example of ambiguous figures, showing competing organization without involving semantic meaning.

explanation. In 1990, Rock and Palmer revealed two primary laws of perceptual organization: “Connectedness” and “common region”, referring to the powerful tendency of the visual system to perceive any connected or enclosed region as a single unit. The phenomena relating to the two laws are so common and self-evident that even classical Gestalt psychologists failed to realize that an explanation was required for why elements that are either physically connected or enclosed by a closed curve are perceived as a single unit. As our discussion

goes on, we will see that these two Gestalt laws closely approach the precise formal (topological) description of the essence of Gestalt principles. Nevertheless, they were neglected for more than a half a century!

Besides the problem of being easily overlooked, one more major challenge to establishing a scientific framework for perceptual organization stems from the abstract and general characteristics of the concept. A theoretical explanation of perceptual organization, to possess explanatory power, must be built on even more general and abstract concepts than this vocabulary. The next question, therefore, is: What kinds of concept are more general and abstract than, for example, “what goes with what”, grouping, belongingness, wholes, and perceptual objects, and therefore, suitable for the formal analysis of organizational processes? It is not difficult to see that featural properties commonly used in the feature-analysis approach, such as orientation, distance, and size, cannot help out in dealing with the problems facing us in finding a formal explanation of perceptual organization.

*Topology provides a formal description of perceptual organization: Insight from invariants over shape-changing transformations*

Topology has been considered a promising mathematical tool for providing a formal analysis of concepts and processing of perceptual organization (e.g., Chen, 1982, 2001). Topology is a branch of mathematics that aims at studying invariant properties and relationships under continuous and one-to-one transformations, termed topological transformations. The properties preserved under an arbitrary topological transformation are called topological properties. A topological framework of visual perception can be broad enough to encompass the variety of phenomena in perceptual organization, such as “what goes with what”, grouping, belonging, and parsing visual scenes into potential objects, and, on the other hand, precise enough to be free from intuitive approaches.

Topology is often considered as one of the most abstract branches of mathematics. If the concepts of topology, their relevance and applicability to perceptual organization are difficult to contemplate in the abstract, an appeal to illustrative examples might be helpful. In the following, I will analyse in some depth three typical cases of perceptual organization to demonstrate why and how to advance the topological approach to perceptual organization.

*Question 1: Figure and ground perception—what attributes of stimuli determine the segregation of figure from background?* Despite the common acceptance that figure–ground perception is fundamental and occurs at the early stage of perception, and despite the large body of empirical findings about

figure-ground phenomena, relatively little literature explicitly discusses precisely what attributes of stimuli determine figure and ground. Challenges to the study of figure-ground perception typically reflect this difficulty in the study of perceptual organization, in being difficult to go beyond intuitive approaches. A mere list of Gestalt laws without unifying formal analysis is unsatisfactory. Is there a common principle underlying various phenomena of figure and ground perception?

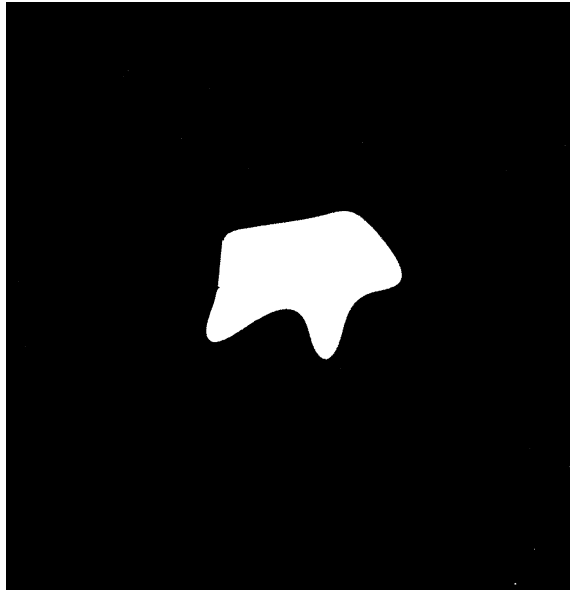
A classical observation reported (Flavell & Draguns, 1957) is that, when a stimulus display is presented under impoverished visual conditions, figure and background achieve some measure of differentiation, although the detailed structure of the stimulus remains vague and amorphous. That the precise shape could not be perceived means that the discrimination of figure from background does not depend on the details of featural properties, such as orientation, location, and size. In other words, a stimulus was separated into different global wholes (a figure and a ground), dependent only on some kinds of global properties, which remain invariant under local changes of featural properties. It appears to be puzzling that these kinds of global properties do exist and lead to some measure of differentiation of figure from ground, but they are shapeless! Exactly what are these shapeless global properties?

To explain the point further, let us examine the Gestalt determinant of “surroundedness” for figure and ground segregation, which is illustrated in Figure 2: A common phenomenon of figure and ground organization is that a surrounded region is likely to be seen as a figure, while the corresponding surrounding region, ground. It is interesting to notice that, as the irregular shape of and the random location of the surrounded region in Figure 2 indicate, the relationship of surroundedness seems to have nothing to do directly with local featural properties like shape, size, location, orientation, and others. Nevertheless, exactly what does “surroundedness” mean? To answer the question, let us appeal to a thought experiment with the display showing “surroundedness”.

**A thought experiment on “surroundedness”.** To appreciate the implication of surroundedness, let us make imagined transformations of the surrounded region in Figure 2. To transform the surrounded region, for example, to move it within (if not outside) the black area, to rotate it, to change its size and even alter its shape. All of these transformations would not cause any difference in perceiving the relationship of the surrounded region as a figure to the surrounding region as a background. In other words, the figure-ground organization based on surroundedness remains invariant under such shape-changing transformations.

This analysis provides an insight into the global nature of attributes determining the relationship of figure and ground: These attributes seem to have nothing to do directly with featural properties, but remain invariant over changes of local featural properties.





**Figure 2.** An illustration of organizational principle of surroundedness in the figure–ground perception.

*Question 2: Apparent motion—what are correspondence tokens in apparent motion with shape-changing transformations?* It is commonly accepted that at the core of understanding apparent motion lies in the correspondence problem. That is, in the process of perceiving apparent motion one has to establish, at some level, a correspondence that identifies which parts or attributes of stimuli presented successively in different frames represent the same object. With respect to the fundamental question of where visual processing begins, the starting question in the study of apparent motion is: What are the constituents of stimuli that are matched by correspondence processes? In his well-known theory of apparent motion, limited to an early feature-analysis approach, Ullman (1979) concluded that while some simple components, such as edge and line segments, are taken to be correspondence tokens matched in perceiving apparent motion, there is no indication that structural forms are part of the correspondence tokens. This conclusion continues to influence research in the area.

**Shape-changing transformations.** One distinguishing aspect of apparent motion is that when one perceives apparent motion, one perceives not only translation and rotation of rigid shapes but also intriguing “plastic deformations”, occurring when apparent motion is produced by dissimilar pairs (Kolars & Pomerantz, 1971). For example, a square moves *and* changes its shape simultaneously to become a triangle or vice versa.

This phenomenon of shape-changing transformations raises the following question: What kinds of invariants still remain and are matched by the correspondence process in such shape-changing transformations? With the transformations of translation and rotation of a rigid shape, simple components like line segments may still keep their identity and, thus, be plausible to serve as correspondence tokens. With shape-changing transformations, it is, however, difficult to imagine how simple objects like line segments act as correspondence tokens. Because, as a shape changes (e.g., a triangle becomes a circle) line segments making up the shape lose their identity, and hence their qualification for being correspondence tokens. This forces us to understand the concept of correspondence tokens in terms of invariants that survive form deformations.

Again, we are confronted with a similar problem to that of “shapeless global properties” discussed in figure-ground perception. A crucial step needed for establishing a formal analysis of the correspondence problem in apparent motion is to describe precisely the holistic identity of object preservation under shape-changing transformations.

*Question 3: The concept of perceptual object, a central concept in a theory of selective attention—how to define the concept of object precisely and formally?* The question “Where visual processing begins?” is embodied in the study of visual selective attention by the question “What does selective attention select?” The majority of theories adopt the basic position of early feature analysis, and assume that attention selects either some of featural properties of stimuli, their spatial location, or their temporal properties. From the viewpoint of early feature analysis, such accounts are natural, because featural properties are held to be primitives and because space and time, as the media in which vision operates, are commonly considered intrinsic to visual processing. An alternative object-based approach to selective attention, however, holds that selective attention selects “objects” or “perceptual units”, which are organized by preliminary perceptual processes (e.g., Desimone & Duncan, 1995; Duncan, 1984; Kahneman & Henik, 1981). The object-based theory claims that to the degree that a perceptual object is attended, all the responses associated with the properties or elements of the object will be facilitated. Thus, the concept of perceptual objects and the rules of perceptual organization that produce perceptual objects should be essential to describe the abilities and limitations of selective attention.

However, the object-based theory faces a main problem, that is, how to formulate precisely the concept of a perceptual object. The concept of a perceptual object is inherited from the Gestalt tradition; like many other Gestalt concepts, it is intuitive, but elusive. Until the intuitive notion of perceptual object is given a precise and formal definition, the claim that selective attention selects objects will not be entirely testable and is in danger of being circular.

Critical to establishing a scientific theory of selective attention is to provide a proper theoretical treatment on the notion of “perceptual object” free from intuitive approaches.

**A thought experiment on perceptual objects.** Since it seems hard to grasp the notion of perceptual objects in the abstract and general sense, let us once again appeal to another thought experiment put forward by Kahneman and Henik (1981):

To appreciate the strength of this notion, imagine a collection of random shapes of various colours, scattered in space. Now imagine that each of these shapes are set in motion, and subjected to gradual distortions and changes of hue. Such a scene will be described as a set of objects that change as they move. The phenomenal impression will be that each object retains its identity over reasonable smooth transformations of any of its properties. There is an obvious distinction between changes and transformations that are identity preserving and others that destroy the original object.

As Kahneman and Henik continue to argue, this primitive notion of identity preservation over transformations is in the core to understand the intuitive notion of perceptual object. This thought experiment indicates that a proper formal analysis of the concept of perceptual object turns out to be an issue of how to give a proper formal definition of “reasonable smooth transformations” and “identity preserving over the reasonable smooth transformations”. The phenomenal impression of identity preservation over shape-changing transformations could provide a starting point for a formal analysis of the intuition of a perceptual object.

*The hypothesis of topological perception: Insight from invariants over shape-changing transformations*

I have considered issues concerned with figure–ground perception, correspondence tokens in apparent motion, and perceptual objects in selective attention. Even though these topics clearly differ in various aspects, in the final analysis, the phenomenological observations associated with each topic are intrinsically related to the concepts of transformation and invariant properties over transformations. Even if figure–ground perception and perceptual objects appear to be divorced from the issue of invariance over transformations, it is necessary to understand their nature from the perspective of identity preservation over shape-changing transformations.

*The central question and a strong conclusion.* Thus, the central question for a formal framework for perceptual organization becomes: Which branch of mathematics is good at describing the unified principle underlying various

expressions of transformations and invariants over transformations, particularly over shape-changing transformations? The previous analyses about invariants over shape-changing transformations lead to a strong conclusion that, for describing principles of perceptual organization, some common mathematical concepts, such as distance and metrics, seem not to be relevant. This implies that any mathematics constructed on the basis of distance and metrics, such as calculus and Euclidean geometry, are not good candidates for a formal system suitable for describing perceptual organization. It follows that in order to establish a formal analysis of perceptual organization, we must look for a different mathematical tool from those commonly applicable in the early feature-analysis approach. My proposition is that topology is just the mathematical language we need.

*Topology and the hypothesis of topological perception.* A topological transformation is, in mathematical terminology, a one-to-one and continuous transformation. Intuitively, it can be imagined as an arbitrary “rubber-sheet” distortion, in which neither breaks nor fusions can happen, however changed in shape the “rubber-sheet” may be. Under this kind of “rubber-sheet” distortion, for example, connectivity, and the number of holes, and the inside/outside relationship remain invariant. Hence they are topological properties. In contrast, local geometric properties such as symmetry, orientation, size, parallelism, and collinearity are not topological properties, because they may be altered by arbitrary “rubber-sheet” distortions. From the perspective of topological analysis, although phenomenally they look quite different, solid figures (e.g., a solid triangle and a disk) are all equivalent to each other, since any one of them can be modified to match any other by performing “rubber-sheet” distortion. On the other hand, “rubber-sheet” distortions cannot create or destroy holes, so a ring containing a hole and a disk containing none are topologically different.

The hypothesis of topological perception assumes that the shape-changing transformations observed in phenomenal world may be precisely described as the topological transformation, and the invariant attributes of an object over shape-changing transformations projected in retinas may be described as three kinds of topological properties in two-dimensional manifolds: Connectivity, the number of holes, and the inside/outside relationship.

Now let us briefly return to the three starting questions discussed above to illustrate why and how a topological approach is applicable to describing perceptual organization. Take the question of “What attributes of stimuli determine the segregation of figure from background?” As I have discussed, the relationship between figure and background depends on global properties that remain invariant over detailed changes of visual features. These global properties may be described as topological in nature, and include connectivity, the number of holes, and inside/outside relations (e.g., Chen, 1982, 2001).



Though originating from separate line of work, Zhang and Wu (1990) and Zhang (1995) proposed a differential-geometric framework for describing neural processes mediating the segregation of figure and ground, in which the topological properties of connectivity and closure of a region are of central concern. As we speak of “an object” in a picture, we usually imply that it is connected (Rosenfeld & Kak, 1982). Similarly, the Gestalt determinant of surroundedness for figure-ground organization is just, in mathematical language, the topological property of holes.

Second, with respect to apparent motion, the topological analysis holds that topological transformations can provide a formal description of shape-changed transformations observed in apparent motion. In particular, topological properties may be considered candidates for correspondence tokens in apparent motion. Take as an example a person walking. Regardless of any changes in featural properties (location, orientation, size, and even shape), the fundamental percept of a segregated object remains; in the language of topology it stays as a single connected component, invariant across feature changes (Chen, 1983). If this percept indeed represents the expression of topological perception, one might expect other topological properties, such as holes, to be good candidates for correspondence tokens in shape-changing apparent motion (Chen, 1985). More generally, shape-changing motion introduces a fundamental factor in visual perception, that is, in Marr’s term (1982, pp. 203–204), “the consistency of an object’s identity through time”. For rigid bodies, the motion correspondence process in time domain, and stereo correspondence process in space domain, are essentially equivalent to each other. Thus, a computational approach, applicable to the correspondence problem in stereopsis, seems to be suitable for the motion correspondence process. However, as Marr realized, for shape-changing bodies, a new theory may be needed for motion correspondence process, because, in shape-changing motion, time introduces a new fundamental issue. This issue is not part of structure-from-motion problem because the precise details of an object’s structure are not relevant. Nevertheless, despite the object-identity problem being at the core of understanding of correspondence problem in particular, and of perceptual organization in general, Marr did not describe any hint as to what this new theory might be. The topological approach, however, provides a new theory for the correspondence problem in shape-changing motion: An object’s holistic identity, preserved under shape-changing motion through time, may be understood in terms of topological concepts, such as holes and connected components, and the percept of maintaining the object’s identity through time may represent an expression of a common underlying principle of topological perception.

Finally, consider the concept of perceptual objects in object-based theories of selective attention. The phenomenal impression of the “reasonable smooth transformations”, highlighted in Kahneman and Henik’s thought experiment

(1981), may be formally described in terms of topological transformations. The meaning of “reasonable” and “smooth” may correspond to “without break and fusion” in the metaphor of rubber-sheet geometry. Identity preservation (as opposed to feature preservation) over transformations—the core intuitive notion of perceptual objects—may be precisely characterized by topological invariants, such as connectivity or connected components. “The distinction between perceptual objects and their properties” (or the distinction between identity preserving and feature preserving) can be considered as the distinction between global topological perception and local featural perception. Kahneman and Henik’s (1981) insight of “the object is prior to its properties and independent of them” may be understood in the following sense that the topological perception is prior to perception of local featural properties, and all of local (metric) properties must be built upon global topological properties (e.g., Chen, 2001; Lappin, 1985). In the following sections, the claim at the core of the topological approach that topological perception is prior to perception of local featural properties will be repeatedly discussed. The “prior” has two strict meanings: First, it implies that global spatial and temporal organizations, determined by topology, are the basis that perception of local geometrical properties depends on; and second, topological perception takes place *earlier* than the perception of local geometrical properties.

## PART II: HOLES AND WHOLE—EVIDENCE SUPPORTING EARLY TOPOLOGICAL PERCEPTION

The hypothesis of *early* topological perception has been tested against a fairly large set of experimental results collected within a variety of paradigms in the study of perceptual organization. The measures and paradigms include, for example: Visual sensitivity (e.g., Chen, 1982), apparent motion (e.g., Chen, 1985; Zhuo et al., 2003), illusory conjunctions (Chen & Zhou, 1997), configural superiority effects (e.g., Todd, Chen, & Norman, 1998), global precedence (e.g., Han, Humphreys, & Chen, 1999a), neuropsychological studies of extinction (Humphreys, Romani, Olson, Riddoch, & Duncan, 1994; Humphreys, 2001), and, pattern discrimination in insects (honeybees) (Chen, Zhang, & Srinivasan, 2003). The results have consistently supported the topological hypothesis.

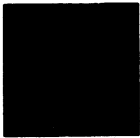
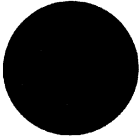



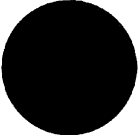
If topological properties play a fundamental role in perceptual organization, we should predict some experimental results that are not necessarily consistent with our everyday perceptual experiences, but with topology. On the one hand, the number of holes mathematically represents a typical kind of topological invariants. On the other hand, our everyday knowledge does not tell us the straightforward psychological implication of holes in perceptual organization. The study of the psychological relevance of holes in perceptual organization, therefore, may serve a fundamental test, going beyond intuitive approaches, for the hypothesis of topological perception.

A major challenge to the study of *early* topological discrimination stems from the fact that there can be, in principle, no two geometric figures that differ only in topological properties, without any differences in nontopological factors. It is, therefore, important to vary approaches and paradigms as widely as possible to test the topological hypothesis from converging operations.

### Visual sensitivity to holes

*The ring vs. the disk.* One of our primary experiments (Chen, 1982) that revealed the robust and counterintuitive influence of topological properties on visual discriminability employed the stimulus displays shown in Figure 3.

The stimulus consists of figures of a triangle, a square, a disk, and a ring, which are commonly considered basic geometrical figures. On the other hand, whereas the solid triangle and the solid square are topologically equivalent to the disk, they are topologically different from the ring that contains a hole. Thus, these pairs of stimulus figures typically represent both local geometrical and global topological distinctions. These figure pairs were also matched in terms of various non-

		% correct response	Area difference (mm <sup>2</sup> )
		43.5	220
		38.5	292
		64.5	254





**Figure 3.** The three stimulus displays used in measuring visual sensitivity to holes in the case of the disk vs. the ring. Also shown are the mean correct response and the area difference for each stimulus pair.

topological factors: For example, luminance flux (the difference between the ring and the disk is larger than that between the solid triangle and the disk but less than that between the solid square and the disk), and sloped edges and curvature (there was less difference between the ring and the disk than between the solid triangle and the disk, or between the solid square and the disk). Furthermore, whereas the triangle and the square appear subjectively to be very different from the disk, the disk and the ring appear more similar to each other.

Subjects received a 5 ms T-scope presentation of one of the three pairs of stimuli, followed by the immediate reappearance of the pre-exposure field, and were asked to report whether the two figures in one display were the same or different (rather than what the two figures are). The intensity of illumination of the target field was adjusted to keep an overall probability of reporting “different” of about 50%. As shown in Figure 3, performance was significantly better for the ring-disk pair than for the other two. Although an intuitive notion of similarity might lead one to expect the opposite result, topology provides an explanation of the data. It is also worth noting that the discriminability between the topologically equivalent figures (the disk vs. the triangle or the square), though intuitively different, were at about the same level. Topology explains all of the results in a unified manner.

*The S-like figure vs. the ring.* An S-like figure vs. a ring shown in Figure 4 were designed to control for luminous flux, spatial frequency components, and other possible confounds of local features. The S-like figure was scaled to approximate the area of the ring, and its shape was purposely made irregular in order to eliminate the possible effects of subjective contours or other organizational factors (such as parallelism, or similarity of length). As a consequence, the S-like figure and the ring differ in holes, but are quite similar in local features, such as luminous flux, spatial frequency components, perimeter length, and averaged edge crossings, in comparison to other two pairs of figures (Figure 4). Despite this, the S-like figure and the ring were most discriminable, and the other stimulus pairs did not differ significantly in discriminability.

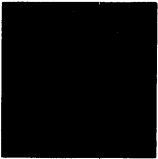
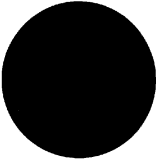
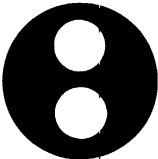
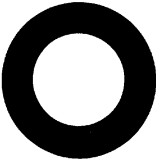
*One hole vs. two holes.* The preceding experiments tested the difference between no holes and one hole, which is just one special case of the more general topological invariant of the number of holes. Tests of the more general case included comparisons of differences in discriminability between shapes with one and two holes (e.g., the ring and the disk with two holes; see Figure 5). The sum of the areas of the two small holes was the same as that of the large hole in the ring. Again high discriminability was found for the topologically different figures, compared with figures that differ in shapes (a solid disk vs. a solid square).

		% correct response	Area difference (mm <sup>2</sup> )
		35.6	310
		59.0	0
		41.1	314

**Figure 4.** In conditions under which various factors commonly considered in the study of visual perception (e.g., luminous flux, spatial frequency components, perimeter length, averaged edge crossings, and other organizational factors) have been controlled, two visual stimuli that are topologically different with respect to holes, such as an S-like figure and a ring, are more discriminable in a near-threshold same–different task than are other pairs of figures that are topologically equivalent, such as a disk and a square or a disk and an S-like figure.

Topological invariants and correspondence tokens  
in apparent motion

As discussed in Part I, topological invariants may play a role in determining apparent motion. Chen (1985) used a competing motion technique to explore the topological structure in apparent motion. Two stimulus displays were successively presented. The first with a single figure in the centre, and the second with two figures located on either side of the centre. For each presentation the subjects were required to choose one of two responses: Motion from the middle figure to the figure on the right or motion to the figure on the left. A series of pairs of displays (Figure 6) were designed to manipulate topological variations between stimuli, while controlling for local features, such as brightness, spatial frequency components,

		% correct response	Area difference (mm <sup>2</sup> )
		39.0	310
		66.2	0

**Figure 5.** The three stimulus displays used in measuring visual sensitivity to holes in the case of the ring and the disk with two holes. Also shown are the mean correct response and the area difference for each stimulus pair (the sum of the areas of the two small holes in the disk was the same as that of the large hole in the ring).

terminators, etc. Subjects displayed a strong preference for motion from a central figure to a topologically equivalent figure, as shown in Table 1. The results demonstrated that topological invariants are indeed used as correspondence tokens in apparent motion.

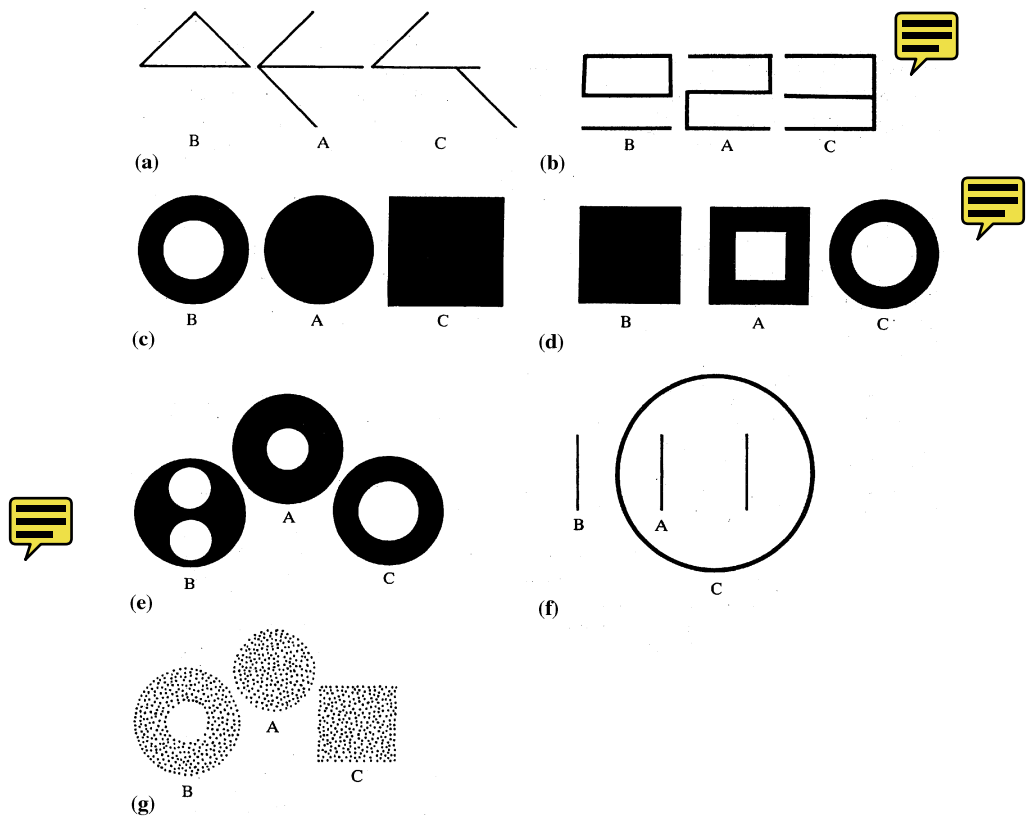
Let us discuss in turn the seven sets of stimuli shown in Figure 6.



- (a) Both the two figures contained in the second stimulus display are made up of exactly the same three line segments as the arrow in the first display. However, the arrow is topologically different from the triangle (in closure<sup>1</sup> or a hole).
- (b) Stimulus A and B were adapted from the stimuli used by Julesz (1981) to test the effect of the number of “terminators” on perception. All of three figures were made up of the same five line segments. Although stimulus A has different number of “terminators” (two) from stimulus C (three) and the same number of terminators as stimulus B, stimulus A (simple connectivity) is topologically different from stimulus B (disconnected, with a hole) but topologically equivalent to stimulus C (simple connectivity).


<sup>1</sup> Closure is essentially a one-dimensional concept. But no real one-dimensional figures can be actually drawn. Rather all drawing figures are two-dimensional. Strictly speaking, the commonly called “closure” is, therefore, “hole” in two-dimensional forms.

- (c) Phenomenally the solid square and disk look quite different but are topologically equivalent, whereas the disk and ring are not topologically equivalent.
- (d) A black square with a square hole replaces the solid black disk in (c). The hollow square and the ring are topologically equivalent, despite that they differ in local features. The stimuli in (c) served to test the effect of topological equivalence of “no holes”. The stimuli in (d) served to test further the effect of topological equivalence of “one hole”. Also, another purpose of (d) was to control for the low spatial-frequency component between the stimuli, bearing in mind the claim that “low



**Figure 6.** The seven experimental displays used to test the topological preference in apparent motion, using a competing motion technique. They were designed to manipulate topological variations between stimuli, while controlling for local features, such as luminous flux, spatial frequency components, terminators, and others. In the procedure of competing motion technique, stimulus A is contained in the first display, stimulus B and C in the second display, in all the seven cases. (See text for details.)

TABLE 1  
Percentages of reports of motion from each middle figure to a  
figure that has the same topological invariants



<i>Item pair</i>	<i>Subjects</i>				<i>Mean</i>
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	
a	92	83	75	75	81
b	83	92	83	83	85
c	83	92	92	83	88
d	92	83	92	92	90
e	83	100	83	92	90
f	92	92	83	83	88
g	83	92	83	75	83

The differences in average percentages between topologically equivalent and not equivalent figures are all significant,  $p < .001$ .

spatial frequencies dominate apparent motion'' (Ramachandran, Ginsburg, & Anstis, 1983). The topological equivalent figures of stimuli A and C were designed to have a similar high spatial-frequency component, whereas the topological equivalent figures in (c), a similar low spatial-frequency components.

- (e) This set of stimuli also served two purposes. First, its primary goal was to test a more general version of the topological hypothesis, that is, the discrimination is sensitive to the number of holes, by varying the difference in the number of holes: One hole vs. one hole and one hole vs. two holes. This manipulation of the number of holes also provide a opportunity to control for the luminous flux factor: The total area of the two inner circles was made to be equal to that of the single inner circle.
- (f) The stimulus figures were used to test the effect of inside/outside relationship, which is another kind of topological invariant, in comparison with distance, which is considered to be critical in computational analysis (e.g., Ullman, 1979). The relation that whether a target figure is within a closed curve or outside it, like the number of holes, also remains invariant under smooth deformations. The central figure (stimulus A) is a vertical line segment. Stimulus B is the same line segment as stimulus A except that it is located at one side of the centre. Stimulus C is made up of a line segment and a circle enclosing the target line. As Figure 6f shows, if the two displays were superimposed, stimulus A would lie within the circle, similarly to the target line of stimulus C, while the target line of stimulus B, outside the circle.



- (g) The figures made up of separate black dots are adapted from Zeeman (1962). They globally look like the stimulus figures in (c): One has a holistic perception, that is, a solid circle in the centre, and a solid square and a ring located on either side of the centre. This holistic and continuous organization suggests that the visual system can ignore details within certain range, attaching importance to global structure based on proximity (Chen, 1982, 2001). This fundamental question will be addressed in detail in Part III. Here we just tested the effect of topological properties based on proximity (rather than point-to-point match) in determining apparent motion.

Regardless of the fact that the seven sets of stimuli represent variations in different topological properties (connectivity, the number of holes, and the inside/outside relationship), and, regardless of the fact that they were designed to control for various nontopological factors (luminous flux, low-spatial frequency components, the number of terminators), the data in all cases supported the topological account of the correspondence tokens in apparent motion: Subjects reported a strong affinity for motion between a central figure and a figure possessing the same topological invariants.

### Holes in illusory conjunctions

Treisman and Gelade (1980), among others, have reported that when attention is overloaded, subjects often report seeing “illusory conjunctions”, that is, incorrect recombinations of the features that are detected. The discovery of illusory conjunctions has stimulated research on numerous issues in the current study of perception and attention, including our own research on the fundamental question “What are the primitives for visual perception?” Specifically, features that give rise to illusory conjunctions must be separately abstracted at an early visual stage, for if such an abstraction of features did not occur, their illusory combination would be difficult to explain. Features that produce illusory conjunctions are therefore strong candidates for the primitives of visual perception.

*Illusory conjunctions of holes.* If topological properties are primitive, illusory conjunctions of topological properties, such as holes, should sometimes take place. To test this prediction, Chen and Zhou (1997) adopted a paradigm originally designed by Treisman (1988). The stimuli each contained three different figures (Figure 7). The subjects’ primary task was to report correctly the two digits flanking on each side of the shapes. As a secondary task, subjects were asked to report the shapes of the three figures in a display, and instructed not to guess but to report only what they were fairly confident of seeing. Each

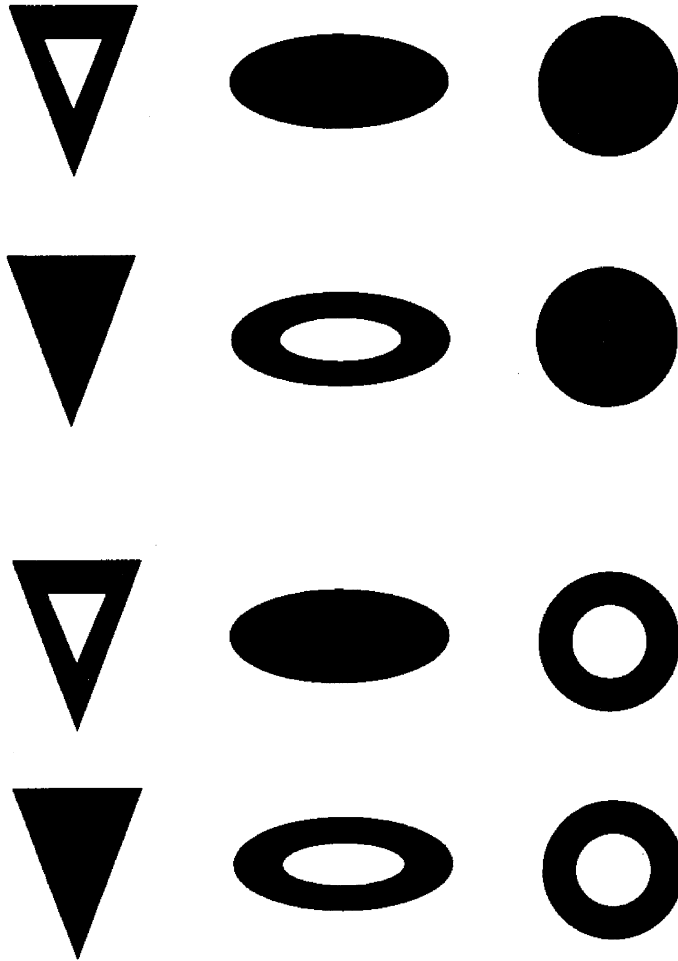


**Figure 7.** (A) An example of the stimulus displays used in the first experiment. Each contained three different figures—a ring, an ellipse, and a triangle—flanked on each side by a digit. (B) An example of the stimulus display, which, as a control for holes, contained all solid figures.

test display was followed by a mask of a checkerboard-like figures. Exposure durations were adjusted separately for each subject to keep an overall probability of correct responses to shapes at a rate of about 90%. Here, correct responses to shapes refer only to outline shapes without considering reports about holes. Errors such as reports of hollow triangle or/and hollow ellipses did not affect the adjustment of exposure durations, for these errors may be due to illusory conjunctions of holes, the occurrence of which was our primary interest.

The results indeed support the prediction of illusory conjunctions of holes; 15 of 19 subjects (16.8% of all trials) reported seeing hollow triangles or/and hollow ellipses (Figure 8), which were not presented. The topological property of a hole contained by a ring appeared to be abstracted, and then recombined with a solid triangle or/and ellipse so that hollow triangles or/and hollow ellipses were produced. The top section of Table 2 summarizes the main aspects of the data.

It might be argued that the subjects might simply misperceive (1) a solid figure as hollow one, and (2) a ring as a hollow ellipse or a hollow triangle. The following analysis, leading to the conservative estimate that illusory conjunction of holes occurs on 10.4% of the trials (shown in the top section of Table 2), rules out this objection. The first kind of feature error commonly occurs in the study of illusory conjunctions: Subjects report a feature, such as a hole, that is not presented in the stimulus displays. For the present experiment, the rate of this type of feature error can be measured by using a control for holes, a display shown also in Figure 7, which contained all solid figures. The rate of this kind of feature error was 2.2% (Table 2). The second kind of feature error is specific to



**Figure 8.** A report was counted as an illusory conjunction only if the subject reported seeing a hollow triangle or a hollow ellipse and, at the same time, a solid disk (top two rows), or a hollow triangle and/or a hollow ellipse, and, at the same time, a ring (bottom two rows).

the present experimental stimuli. If subjects misperceive a ring as a hollow ellipse or a hollow triangle, that is, misperceive the outer border of the ring as an ellipse or triangle, what could such reports look like? Subjects may report a hollow ellipse or hollow triangle but no ring or disk, as illustrated in Figure 9. This type of error occurred at most on only 4.2% of the trials (Table 2). Note that these errors could also count as illusory conjunctions (i.e., an illusory conjunction of the hole, along with an error in the perception of the shape). Thus, the excess of conjunction errors over the sum of these two types of feature

TABLE 2  
Mean percentage of correct reports, illusory conjunctions, and feature errors of holes in  
the experiment of reporting by hand drawing

<i>Response</i>	<i>Percentage of total reports</i>
<i>Verbally reporting</i>	
Correct shapes	88.6
Errors of outline shapes	11.4*
Possible illusory conjunctions of holes	16.8
Feature errors of holes:	
1. Misperceiving a solid figure as a hollow one	2.2
2. A. Misperceiving a ring as a hollow triangle	2.3
B. Misperceiving a ring as a hollow ellipse	1.9
Conservative estimate of illusory conjunctions	10.4**
<i>Reporting by hand drawing experiment</i>	
Correct shapes	90.6
Errors of outline shapes	9.4
Possible illusory conjunctions of holes	24.0
Feature errors of holes:	
1. Misperceiving a solid figure as a hollow one	1.8
2. A. Reporting a hollow triangle but without a ring or disk	2.2
B. Reporting a hollow ellipse but without a ring or disk	2.4
Conservative estimate of illusory conjunctions	17.6***

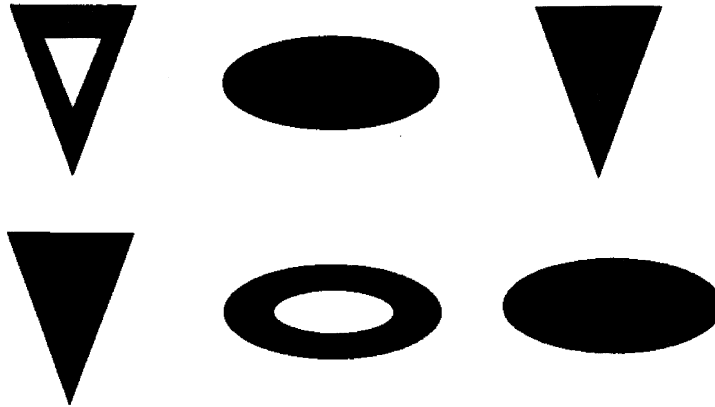
\* Under the condition of brief presentation with masking and divided attention, subjects made errors in reporting spontaneously various kinds of outline shapes, which were not presented, for example, squares, rectangles, upward triangles or ellipses, repeated shapes like two disks or two ellipses or two triangles at one presentation, and other irregular shapes.

\*\* The difference between illusory conjunctions and feature errors was significantly greater than zero for both the feature errors of type (1) ( $t$ -test,  $p < .001$ ), and type (2), ( $t$ -test,  $p < .001$ ), and for the sum of the two types of feature errors ( $t$ -test,  $p < .002$ ).

\*\*\* The difference between illusory conjunctions and feature errors was significantly greater than zero for both the feature errors of type (1) ( $t$ -test,  $p < .001$ ), and type (2), ( $t$ -test,  $p < .001$ ), and for the sum of the two types of feature errors ( $t$ -test,  $p < .001$ ).

errors, 10.4% ( $16.8 - 2.2 - 4.2\%$ ), represents a conservative estimate of illusory conjunctions, rather than simply misperceptions of holes.

*The abstract nature of holes.* To examine further the confidence of the subjects in their reports, they were asked to point out which of the figures they had perceived from among a list of various kinds of figures in the first two and the last two blocks of presentations. The list of figures include, in addition to the four kinds of figures actually presented: Two kinds of hollow triangles—a triangle with a triangle hole and a triangle with a circular hole; two kinds of hollow ellipses—an ellipse with an elliptic hole and an ellipse with a circular



**Figure 9.** Reports of hollow triangles and hollow ellipses other than those in Figure 8 were not counted as illusory conjunctions. Instead, the reports were attributed to Type 2 feature errors of holes (see Table 2).

hole; a solid rectangle; and a solid square. The results further support the argument that the conjunction errors involving holes are genuine perceptual experiences in the following two senses. First, 17 subjects, including 2 who did not verbally report seeing hollow triangles or ellipses, pointed out hollow triangles and/or ellipses at a rate of 21% of the total presentations; but only 3 subjects pointed out solid rectangles and/or squares, with an average rate of less than 1%. Second, it is particularly interesting that only triangles containing a triangle hole and ellipses containing an elliptic hole were reported by pointing; in other words, no subjects reported seeing triangle containing a circular hole and/or an ellipse containing a circular hole. If the subjects simply misperceiving the outline shape of the ring or simply superimposed the circular hole contained by the ring on the solid figures because of, say, afterimages, it would be difficult to explain why no circular holes with triangles or ellipses were perceived. This aspect of the result also indicates the abstract nature of the holes containing in the illusory hollow figures, that is, the concrete shape of a hole is irrelevant to its identity as a hole. The abstract nature suggests that holes were extracted before their illusory conjunctions.

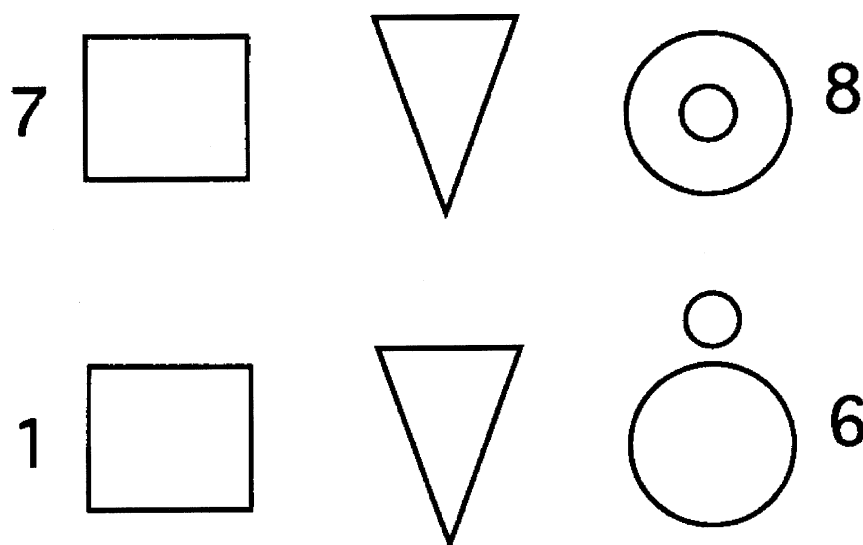
Eight out of nineteen subjects who reported seeing hollow triangles and/or hollow ellipses by pointing said at that time that in their verbal reports they often did not distinguishing between a hollow figure and a solid one of the same outline shape. This aspect of the results called attention to a limitation on the verbal reporting of shapes. On the other hand, to instruct subjects to report particularly if there are holes might introduce the possibility of bias. In view of these considerations, one more experiment further tested the illusory conjunction of holes, using a different procedure for reporting—namely, drawing by hand instead of reporting verbally. As can seen in the bottom half of Table 2, the

results showed illusory conjunctions on 17.6% of the trials (much higher than those shown by verbal report in the previous experiment), and further confirmed the illusory conjunction of holes.

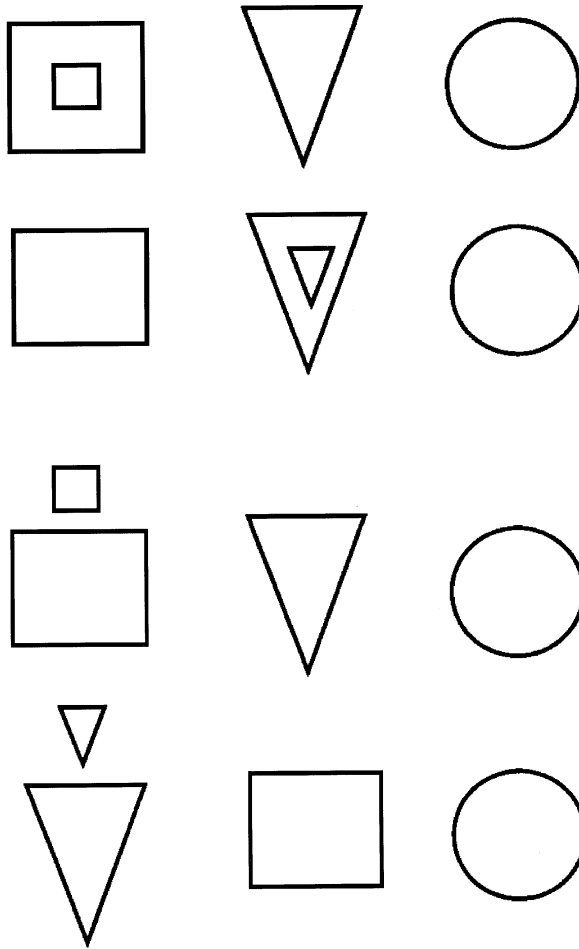
*The constraint of inside/outside relationship in illusory conjunctions.* One distinguishing aspect of illusory hollow figures produced in the experiments is that the conjoined holes underwent remarkable geometric transformations, such as shape transformations of a circle to a triangle or to an ellipse. The next question is which constraints, if any, play a role in such drastic deformations shown by illusory holes. One more experiment was run to examine the constraints. Subjects were shown a set of three outline figures (Figure 10), which represent the inside/outside relationship.

With stimulus A, all subjects reported seeing a small figure contained within a larger triangle or square, and with stimulus B, a small figure outside a larger triangle or square. To keep the estimate of illusory conjunctions conservative, only the response sets shown in Figure 11 were considered to be illusory conjunctions. They occurred on 16.6% ( $t$ -test,  $p < .001$ ) and 12.4% ( $t$ -test,  $p < .002$ ), respectively, of the total presentations.

The following points in relation to Figure 10 are noteworthy: (1) With the stimuli of Figure 10A, the same shape transformation was observed. All subjects except one reported seeing only a small triangle and square, enclosed by a larger triangle and square, respectively, rather than the small circle that was



**Figure 10.** (A) Stimulus displays representing the inside relationship. (B) Stimulus displays used, as a control for the inside/outside relationship, to represent the outside relationship.



**Figure 11.** Examples of the response sets that were considered to be illusory conjunctions. To keep the estimate of illusory conjunction conservative, only these response sets were counted.

actually presented; (2) while with stimulus A no subject reported seeing a small figure outside a larger one, with stimulus B no subject reported a small figure encircled by a larger one either. This indicates that, although transformations of shape occur in the production of illusory conjunctions, two topological constraints, those of number of holes and the inside/outside relationship, remain invariant.

*The implication of the abstract nature of holes independent of featural properties.* As pointed out before, a major difficulty with the evidence for

early topological perception is that it is difficult to design stimuli that represent only topological distinctions without introducing differences in nontopological variables. Even though factors may be controlled under different conditions, any single piece of evidence for topological perception still leaves room for an argument based on confounding by nontopological factors. The conjoined holes observed in illusory conjunctions, however, underwent shape-changing transformations. This indicates that the holes were perceived as abstract entities available at an early stage of vision, independent of local geometric features. This confounding is, therefore, unlikely to account for illusory conjunctions with holes, which differ from the original ones in shape.

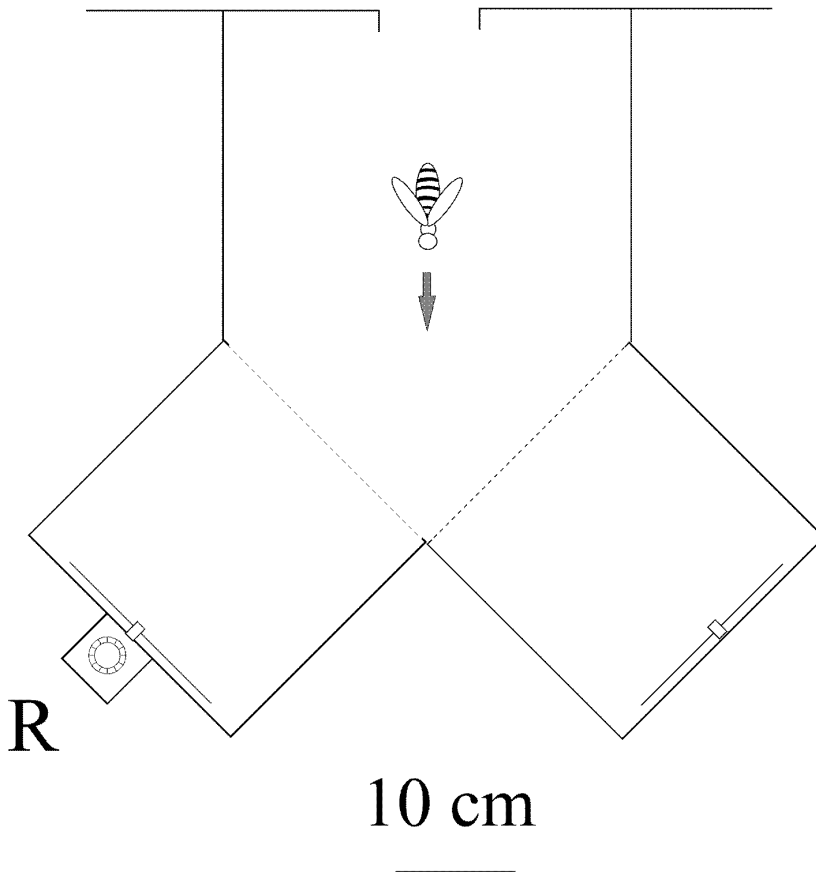
### Global perception in small brains: Topological pattern recognition in honeybees

In human visual perception, early pattern recognition engages later, cognitive, stages, which may make the primitives of perception more difficult to discern. Studies of pattern recognition in small creatures like honeybees with their relatively simple nervous systems might reveal certain visual capacities that are shared across species, which may therefore represent elementary and general underlying functions for any visual system (Srinivasan & Zhang, 1998). If topological perception is indeed a fundamental property of vision, one might expect topological properties to be extracted by all visual systems, including the relatively simple ones possessed by insects. Thus, a rudimentary visual system may not be good at discriminating patterns that are topologically equivalent but be able to discriminate patterns that are topologically different. A series of experiments with bees demonstrate that their small brains nevertheless possess the ability for topological perception.

A Y-maze paradigm (Wehner, 1971; Figure 12) was used to train bees to distinguish between patterns. In each experiment, a fresh group of 8–10 bees was marked and trained to enter the Y-maze apparatus that presented two stimuli, one on the vertical end wall of each tunnel. One stimulus (termed “positive”) offered a reward of sugar water, R, which the bees could reach through a tube. The other stimulus (termed “negative”) carried no reward. The positions of the positive and negative stimuli were interchanged every 10 minutes.

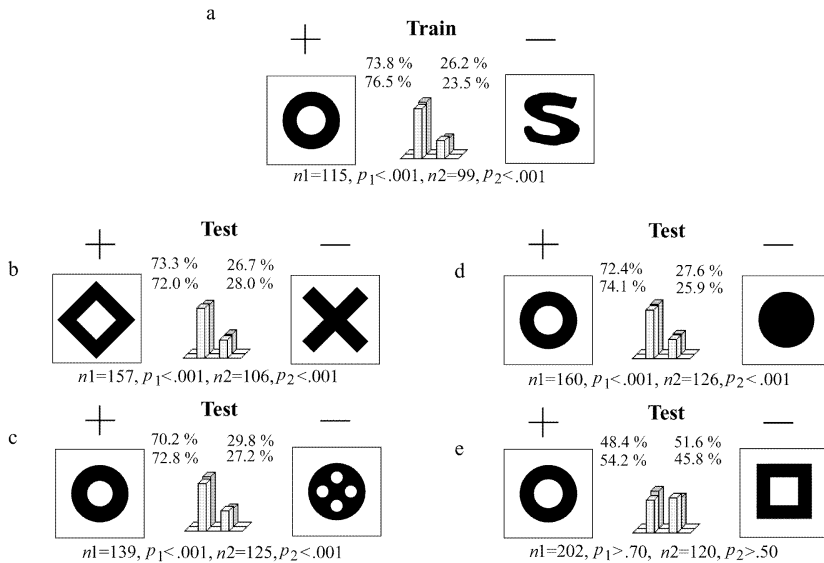
Experiment 1 investigated topologically based discrimination using five pairs of stimuli, shown in Figure 13a–e. The patterns were designed to present topological differences, and to exclude the use of nontopological cues in making the discriminations. They were: A ring and an S (Figure 13a); a hollow diamond and a cross-like figure (hereafter referred to as a cross) (Figure 13b); a ring and a disk with four holes (Figure 13c); a ring and a solid disk (3) (Figure 13d); and a ring and a hollow square (Figure 13e). The two figures in each of the first four pairs (a–d) differ topologically in the number of holes. The ring and the hollow





**Figure 12.** Y-maze apparatus for training and testing bees. Scale applies to the apparatus.

diamond each contain a hole while the S, the cross and the disk do not, and the ring and the four-hole disk also differ in the number of holes they contain. At the same time, these stimulus pairs were designed to control carefully for non-topological features. In Pair a, the ring and S were made to have equal area (and therefore luminous flux), very nearly the same spatial frequency components and perimeter length, and equal averaged edge crossings. The shape of the S was also made to be irregular to eliminate possible effects of subjective contours. In Pair b, the hollow diamond and the cross were oriented with their edges parallel to eliminate potential use of orientation cues, and also made to have equal area. In Pair c, the total area of the four smaller holes contained within the four-hole disk was made equal to the area of the larger hole contained within the ring, and their spatial-frequency spectrum and perimeter length were nearly equal.



**Figure 13.** (a) Training stimulus pair. (b–e) Test stimulus pairs. Bars and numbers show relative frequencies of choices in favour of the positive (+) and negative (–) stimuli, as measured after training and in various transfer tests. Light and dark bars depict results from Experiments 1 and 2, respectively.  $n1$  and  $n2$  are the number of choices analysed in Experiments 1 and 2 respectively, and  $p1$  and  $p2$  are the respective values associated with a  $\chi^2$  test for significant difference from random-choice behaviour. In a separate experiment, bees were trained to distinguish between an S and a ring with the S as the positive stimulus. The stimulus figures were all photographic reductions of the original ones, and were all “white on black” forms.

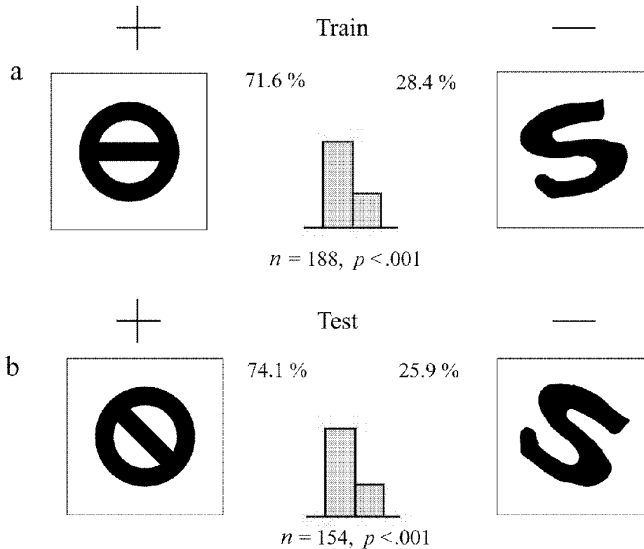
Although phenomenally they look quite different, the ring and hollow square in Pair e are topologically equivalent to each other, since each of them contains the same number of holes, specifically one hole in this case.

Bees were first trained to discriminate between the ring and the S, by rewarding them on the ring. Despite the fact that the ring and the S possess the same area and nearly the same spatial-frequency spectrum, this task was learned well. The choice frequency in favour of the positive stimulus is 73.8% ( $n1 = 115, p < .001$ ). The trained bees were then tested with each of the stimulus pairs b–e, in turn, as shown in Figure 13. The bees were immediately able to distinguish between the patterns presented in the test pairs b, c, and d, without any training on these specific patterns. The choice frequencies in favour of the positive stimulus are 73.8% ( $n1 = 157, p < .001$ ), 70.2% ( $n1 = 139, p < .001$ ), 72.4% ( $n1 = 160, p < .001$ ), respectively. That is, bees that had been trained to distinguish between the ring and the S could immediately distinguish between a hollow diamond and a cross, between a ring and a four-hole disk, and between a

ring and a disk. In other words, the bees behaved as though they perceived and learned the topological difference between the ring and the S and were using this as a cue to distinguish between the other patterns, which they had never previously encountered. This occurred despite the fact that the patterns in each of Pairs b to d differ from those in Pair a with regard to various local features. By contrast, this discrimination capacity did not transfer to the testing stimulus of Figure 13e. That is, the bees were unable to distinguish between the ring and the hollow square, which are topologically equivalent, even though they appear more different than the ring and the disk, and differ in local features such as orientation and curvature. The choice frequency in favour of the positive stimulus was 48.4% ( $n1 = 202$ ,  $p > .70$ ).

The above results were replicated in Experiment 2, in which a Y-maze modified by adding a transparent baffle (with a central hole 5 cm in diameter) at the entrance to each arm, was used. This modified Y-maze forced the bees to slow down and inspect the stimuli more carefully before making a decision and to allow the bees' choices to be monitored more precisely. The results are also shown in Figure 13.

One of the primary concerns in the study of pattern recognition in bees is the role played by local cues in distinguishing visual patterns. Thus, it could be argued that the discrimination capacity observed in at least some of the above experiments may have arisen from differences in local, nontopological cues rather than topological difference. For instance, the ring could have been distinguished from the S on the basis of the fact that the ring carries a white part at the middle (e.g., stimulating an on-centre cell), or the S could have been distinguished from the ring on the basis that the S carries an oriented, straight-line segment in the middle. To address these possibilities, we conducted Experiment 3 in which a fresh group of bees was trained to discriminate between a  $\theta$  and an S (Figure 14a). Neither of these stimuli contains a white region at the middle, but they are topologically different because the  $\theta$  contains two holes whereas the S contains none. The  $\theta$  and the S were made to have the same area. To prevent the possible use of local orientation cues in making the discrimination, the two stimuli were always oriented such that the central line segments of the S and the  $\theta$  were parallel. Furthermore, to prevent the possible use of other local cues such as black areas, the stimuli were both rotated by  $90^\circ$  at regular intervals during the training period. The rotation was performed each time the positions of the positive and negative stimuli were interchanged in the Y-maze. After training, the bees were tested with the same pair of stimuli, but each stimulus was rotated by  $45^\circ$  relative to its orientations during training (Figure 14b). In other words, during the tests, the stimuli were both oriented at  $+45^\circ$  or  $-45^\circ$ , each orientation being presented for half the duration of the test. This procedure for training and testing rendered the use of local intensity and orientation cues very unlikely. Nonetheless, the bees were clearly able to distinguish the  $\theta$  from the S in the training as well as in the tests. The choice

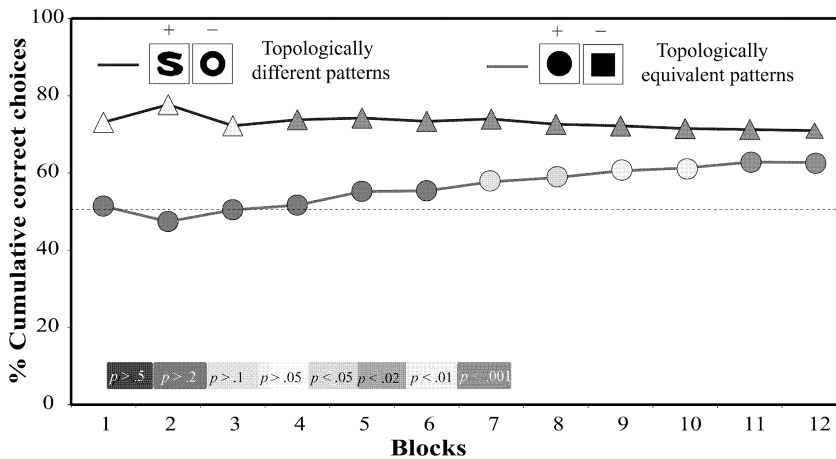


**Figure 14.** (a) Training and (b) test stimulus pairs used in Experiment 3 to examine the possible role of local differences in pattern intensity on discrimination. The S and the  $\Theta$  were photographic reductions of the original ones, and were “white on black” forms. The choice frequencies, the number of choices ( $n$ ), and the  $p$  value in a  $\chi^2$  test for significant differences from random choice are given in the figure. (See text for details.)

frequencies in favour of the positive stimulus of the  $\theta$  are 76.1% ( $n = 188$ ,  $p < .001$ ) in the training duration, and 74.1% ( $n = 154$ ,  $p < .001$ ) in testing duration, respectively.

Measurements of learning curves indicate that bees learn to recognize topological differences very rapidly (Figure 15). One block represents approx. four rewards per bee, on average. Twelve blocks were run for each pair of stimuli to attain the plateaus of the learning curves. When bees are trained to distinguish between the S and the ring, the plateau of the learning curve is attained very quickly. In fact, learning occurs so rapidly that it is difficult trace its development in the initial part of the curve. From block 1 onwards, the preference for the correct stimulus is already significantly different from the random-choice level. In contrast, when trained to distinguish between shapes that are topologically equivalent—a disk and a solid square, despite their apparent differences in local features (such as extended horizontals and verticals), the plateau is attained only after 11 blocks, and performance departs from random choice only after block 7.

The main conclusions from these bee experiments are: (1) Bees are able to abstract the general property of topological invariance, namely, the number of



**Figure 15.** Learning curves for two different experiments. Triangles: Performance in distinguishing between two topologically different stimuli (the S and the ring). Circles: Performance in distinguishing between two topologically equivalent stimuli (the disk and the square). In each case, the choice frequencies are accumulated from the start of the experiment, and averaged over all participating bees. One block represents approx. four rewards per bee, on average. The dashed line depicts the random choice level (50%). Total choices in each block range from 28 to 43.

holes; and (2) bees are capable of learning topologically based discriminations very rapidly. These results of topological discrimination in bees are particularly useful in revealing the topological primitives of visual representation. The holes in the testing stimuli were designed to be different from those in the training stimuli in widely varied local features. The transferability demonstrates the abstract nature of holes, independent of the local features. It is therefore quite difficult to explain the topological discrimination in terms of confounds of such nontopological features. In addition, the learning curves in performance evolution provide a unique criterion for distinguishing topological discrimination from discriminations based on nontopological features. Thus, topological pattern recognition may be a fundamental aspect of bees' visual processing. It is well known that computational theories of vision tend to assume that the primitive elements of computation are local geometrical features (Marr, 1982). With respect to such local primitives, topological properties have high computational complexity (Minsky & Papert, 1969). Existing computational models, therefore, lead one to expect that discriminations based on topological properties would occur at a higher level of perception than those based on local geometrical properties. We have shown here, however, that topological perception is displayed by a creature with a brain weighing less than a tenth of a milligram and carrying fewer than 0.01% as many neurons as the human brain. We also find that bees are quicker at learning discrimination

tasks if the discrimination was based on topological cues. These findings contradict prevailing notions of computational complexity. Although the global nature of topological properties makes their computation difficult, topology may be a fundamental component of the vocabulary of visual systems across species.

### PART III: TOLERANCE SPACES AND PERCEPTUAL ORGANIZATION—GLOBAL PROPERTIES IN A DISCRETE SET AND THEIR APPLICABILITY TO PERCEPTUAL ORGANIZATION

So far it seems that we are quite successful in showing a wide applicability of the topological approach to various issues in perceptual organization. There exists, however, a fundamental obstacle to applying topology to the study of perception, one that concerns the most basic and simple phenomena in perceptual organization, that is, an apparently disconnected array may be perceived as a subjectively connected object. This phenomenon of subject connectedness despite discrete stimulus input seems to contradict the mathematics of topology. Should such kind of Gestalt phenomena be considered as counterexamples against the topological approach? To answer this challenge, we need to examine the very notion of global properties in perceptual organization.

#### How to define global properties in a discrete set

Visual processing is essentially discrete, starting from sampling continuous stimuli by the cone and rod mosaic of the retina. In contrast, our percepts of individual objects and events in the outside world are continuous and holistic. The very notion of perceptual organization concerns the fundamental and ubiquitous phenomena that our perceptual world is composed of mutually separate and yet individually connected perceptual objects. Particularly, the Gestalt laws are often based on apparently disconnected stimuli, such as disconnected dots arrays, and deal with such question as: How are disconnected arrays organized into perceptually connected wholes?

This contradiction strongly indicates that we are not able to directly apply general (continuous) topology or point-set topology to describe perceptual organization. What we need is a special branch of topology, which should be, on the one hand, like general topology, good for describing global properties such as connectivity and holes, and, on the other hand, applicable to discrete sets. A proper mathematical treatment on this issue is a critical step for introducing topology to describe perceptual organization. The above analysis leads to a fundamental question: How to mathematically define global properties in a

discrete set? Or what attributes that belong to a physically disconnected stimulus (such as a dot array) determine the perceptual connectivity?

To understand this, we need to consider the mathematical idea of “tolerance spaces”, introduced by Zeeman (1962). I start with a discussion of the concept of tolerance, which plays a central role in the mathematical structures of tolerance spaces.

### Tolerance, and its psychological relevance

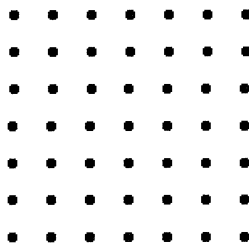
Our perceptual systems have a limited capability for detecting or discriminating stimuli. For example, with respect to visual acuity, two sufficiently close dots in the visual field are not discriminable. Zeeman (1962) advanced a mathematical structure of a “tolerance” as a mathematical abstraction of the nature of psychological “least noticeable difference”, covering topics such as spatial acuity and contrast sensitivity. A tolerance is defined in terms of an algebraic relation, which is reflexive and symmetric, but in general not transitive. In algebraic language, suppose  $X$  is a set of stimuli. If two stimuli  $x_1, x_2$  in  $X$  are sufficiently close so as not to be distinguished we say they are within a tolerance, and the notation  $x_1 \sim x_2$  is used. The tolerance  $\xi$  is defined to be the set of pairs  $(x_1, x_2)$  such that  $x_1 \sim x_2$ . A tolerance  $\xi$  is reflexive. That means, for any element  $x$  belonging to set  $X$ ,  $(x, x)$  belongs to the tolerance  $\xi$ . The physical meaning of “reflexive” is obvious: For example, a dot (or any element in a physical set) is not discriminable from itself. A tolerance  $\xi$  is symmetric. That means, if, for two elements  $x_1$  and  $x_2$  belonging to set  $X$ ,  $(x_1, x_2)$  belongs to the tolerance  $\xi$ , then  $(x_2, x_1)$  also belongs to the tolerance  $\xi$ . It is also easy to see the physical meaning of “symmetric”. For example, in the visual field, if two dots are sufficiently close to each other so that the first dot is not discriminable from the second dot, then the reverse relation is also true that the second dot is not discriminable from the first dot either. It is important to notice that a tolerance is not in general transitive, that is to say if  $(x_1, x_2)$  and  $(x_2, x_3)$  belong to the tolerance  $\xi$  this does not imply that  $(x_1, x_3)$  belongs to the tolerance  $\xi$ . The nontransitivity is basic for the very notion of a tolerance, otherwise all of elements that are covered in sequence by “the coat of mail” of a tolerance will be indistinguishable, and the range of the tolerance will expand—contrary to the meaning of least noticeable difference. The above analysis indicates that the mathematical definition of tolerance serves a proper abstraction of the practical meaning of the least noticeable difference.

However, even though originally the implication of a tolerance used by Zeeman (1962) corresponds to the concept of the least noticeable difference, for the purpose of understanding global properties in a discrete set, the mathematical nature of tolerance actually generates broader implications than that implied by the least noticeable difference. The notion of least noticeable difference refers to being unable to distinguish sufficiently close variables. In

contrast, the notion of tolerance, when applied to perceptual organization, could imply ignoring detailed changes (which may still be distinguishable) for attaching importance to global properties. In this sense, a least noticeable difference is just a special case of a tolerance. What makes this definition of tolerance interesting is that it appears to be a basic strategy for the visual system to perceive global properties: Detailed structures or changes within a tolerance are treated with indifference by the perceptual system in order to capture more global properties. Let us take a very common phenomenon, shown in Figure 16, as an example to illuminate the psychological relevance of tolerance.

A phenomenal response to the dot matrix of Figure 16 is that, while we can clearly see the matrix as a disconnected array made up by separate dots, we can also perceive the matrix as a connected whole as long as two adjacent dots are sufficiently close to each other. The holistic percept of the disconnected matrix indicates a fundamental aspect of perceptual organization, that is, a clearly disconnected array is perceived as a holistic object, as if the visual system ignores “on purpose” detailed gaps within a tolerance for emphasizing the global organization of the stimulus array. Hence, it is important that the notion of a tolerance should not be restricted to the least noticeable difference, such as a resolution or a threshold. On the other hand, a solid line segment is automatically perceived to be connected and continuous. Yet, it is impossible to draw a connected line segment in the mathematical sense. When seen under high resolution, the line segment can appear disconnected. Then, the same issue remains why a line segment is perceived as continuous when it is physically disconnected? Hence the concept of tolerance should serve a consistent mathematical structure suitable for describing perceptual organization of subunits, which are discriminable or not but within a tolerance.

To illustrate further the physical implication of tolerance, let us appeal to the following metaphor proposed by Zeeman (1965): “A tolerance on a set



**Figure 16.** An example to illuminate the psychological relevance of tolerance. The holistic percept of the disconnected matrix indicates a fundamental aspect of perceptual organization, that is, a clearly disconnected array is perceived as a holistic object, as if the visual system ignores “on purpose” detailed gaps within a tolerance for emphasizing the global organization of the stimulus array.



acts as a kind of glue, gluing the set together, rather as a topology glues a topological space together". Consider now, for example, that a set is the atoms in the paper of which this page is made, and the tolerance was one millimetre. "The tolerance is rather like the solid-state forces between the atoms binding them together into the piece of paper. Were the paper made of rubber, we would bend it and stretch it without essentially disturbing the tolerance, provided that we did not cut it or glue it. Therefore global tolerance properties are exactly like global topological properties, and the two most noticeable of these properties are that the paper is all in one piece, and that it is 2-dimensional." (Zeeman, 1965).

*Ecological minimum measure.* One aspect of Gibson's ecological approach to perception (1979) concerns minimum measures in the ecological environment, that is, the ecologically meaningful scale of space and time, which are different from those in the physical world. As emphasized by Gibson, in the space domain, the size-level at which the ecological environment exists is an intermediate (being measured in metres or centimetres), compared with the level of atoms and basic particles (being measured in millionth of a millimetre and less) at one extreme, and to the level of galaxies (measured in light-years and more) at the other. In the time domain, the temporal scale of ecological process and events is also intermediate (being measured in days, hours, and seconds), compared with atomic physics (measured in millionth of a second) at one extreme and astronomy (measured in millions of years) at the other. The minimum measure specific for the ecological environment makes the human perceptual system psychologically specific, and ecologically powerful. Compared with atomic physics and Newton mechanics, the unique laws of statistic physics may be due to its specific minimum measure of space, which is smaller enough than that of Newton mechanics (so that it makes sense to talk about individual particles and their single activities), and at the same time, larger enough than that of atomic physics (so that it makes sense to talk about a group of particles and their statistical behaviour). Logically, the ecological minimum measures of space and time play a fundamental role in the function of human perception, and cause unique characteristics of human perceptual process. Perceptual concepts always, to some extent, depend on certain minimum measures. Obviously, apart from the consideration of minimum measure of discriminability, subjective concepts, such as perceptual "clearness", "straight", "flat", and "smooth" would be meaningless. The above analyses suggest that we need some concept like tolerance to provide a formal description of a minimum measure. A tolerance should be considered as a mathematical abstraction of the implication of a minimum measure, as they both refer to a range within which detailed structures or changes are treated with indifference by the perceptual system for attaching importance to global properties.

*A hierarchy of elementary temporal experiences and minimum measures.* Pöppel (e.g., 1985, 1997) proposed “a hierarchical model of temporal perception”. In the time domain, a systematic manipulation of variables of duration or interval available for processing revealed a hierarchy of “elementary temporal experiences” (ETEs) in temporal organization: From the subjective “simultaneity”, to the subjective “temporal order”, and to the subjective “now”. Below the fusion thresholds of about 2–3 ms, for two objectively nonsimultaneous tones, subjects find themselves inside a single “window of simultaneity”. In the next step, above the fusion threshold but below the order threshold of about 30–40 ms, there is a stage of the “incomplete” simultaneity; subjects can tell, to be sure, that two stimuli occur nonsimultaneously, but cannot tell the order of them. Only beyond the boundary can subjects say with sufficient assurance that two stimuli were nonsimultaneous because one occurred first and another second. These results suggest the operation of a form of temporal tolerance that determines the experiences of simultaneity and sequentiality (or chronological order), which are not directly derived from absolute physical simultaneity and sequentiality. It is even more interesting to observe that there is a tolerance of timescale (in the neighbourhood of 2.5–3 s) within which apparently separate successive events are treated with indifference for being integrated into one perceptual unit or a temporal Gestalt. But beyond the tolerance of timescale the integration to a single unit then breaks apart. Take the example of hearing metronome beats within up to 2.5–3 s. While we clearly perceive that the sequences of beats are composed of the discontinuous beats, we do have a holistic perception of hearing the unified whole formed by beats. The percept of the sequence of beats here is just like that of the dot array in vision in the sense that in both cases the perceptual system seems to ignore local variations of either space or time in order to capture their global organization. The very notion of “now” or “present” depends on an integrative mechanism that fuses chronologically sequential events into “a present Gestalt”. As Pöppel (1985) argued, “If we now read or hear the word ‘now’, we read or hear the whole ‘now’ now. We do not read or hear the sequence of three different letters or speech sounds, n-o-w. Apparently, the sequence of letters is fused in our experience into a unit of perception.” When we speculate about time on the basis of Newton’s conception that time flows continuously and can be recorded with unlimited accuracy if an instrument used is accurate enough, it may be logical to say that the concept of “the now” or “the present” is the nonextensive boundary between past and future. However, to regard “now” only as a nonextensive boundary between past and future fails to correspond to our experience. The participation of ecological temporal tolerances in perception gives rise to our feelings of simultaneity, sequentiality, and present moment, which are fundamentally different from their counterparts in the physical world. Obviously, the perceptual temporal structure of ETSs determined by ecological minimum measures is fundamental constraint on all mental states and processes.

The concept of tolerance may provide a unified mathematical description of minimum measures of both space and time domains in the ecological environment.

In summary, the mathematical abstraction of a tolerance provided a unified basis for describing global properties in a discrete set. A tolerance refers to the range within which detailed variations are ignored for attaching importance to global properties. The tolerance and the global properties of a discrete set, therefore, are two sides of a coin, and are mutually dependent concepts. At first sight a tolerance looks a limitation in perceptual function, but in fact it is one of the most fundamental strengths of the perceptual system, which allows the perceptual system put its source to best use. Tolerance spaces may provide a mathematical structure necessary to meet the challenge posed by the task of establishing general principles for perceptual organization at the scale of ecological minimum measures in both space and temporal domain.

### Mathematics of tolerance spaces: To define global properties in a discrete set

A set  $X$  together with a tolerance  $\xi$  is called a tolerance space  $(X, \xi)$ . Elements in a tolerance space need not be restricted to physical dots in space but may be any variables causing perceptual organization. This abstract characteristic allows tolerance spaces to serve as a unified description of different forms of perceptual groupings, including proximity and similarity. There are more examples of tolerance spaces.

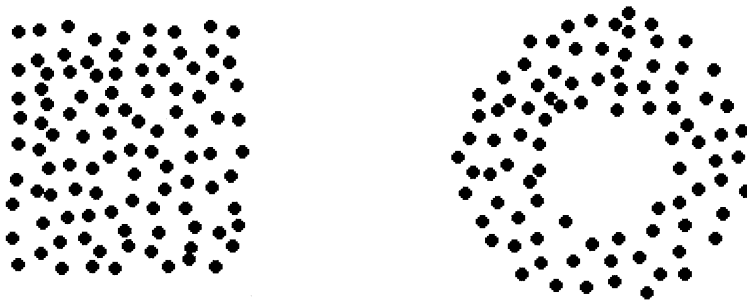
- *Example 1.* Let  $X$  be a finite set of discrete dots arranged in a row. The tolerance  $\xi$  is the set of all pairs of adjacent dots. The set of dots  $X$  together with the tolerance  $\xi$  is a tolerance space denoted as  $(X, \xi)$ .
- *Example 2.* Let  $X$  be the visual field, and  $\xi$  be the visual acuity tolerance, that is to say, all pairs of points that are indistinguishable. The visual field  $X$  together with the visual acuity  $\hat{i}$  is a tolerance space denoted as  $(X, \xi)$ .
- *Example 3.* Let  $X$  be the Euclidean plane, and tolerance  $\xi$  all pairs of points in the plane that are less than a distance of  $\varepsilon$  apart. This tolerance space is denoted as  $(X, \xi)$ .
- *Example 4.* Let  $X$  be the set of contrasts, and  $\xi$  be contrast sensitivity tolerance, that is to say, all pairs of contrasts that are indistinguishable. The set of contrasts  $X$  together with the contrast sensitivity  $\xi$  is a tolerance space  $(X, \xi)$ .

A tolerance is, as emphasized above, not transitive. In comparison with the algebraic relation of equivalence that is reflexive and symmetrical as well as transitive, the special feature of intransitive gives the notion of tolerance spaces

special mathematical properties: Global topological properties are highly emphasized. While mathematical forms such differential equations are appropriate to describe changes of local properties, the global tolerance structure is proper to treat global properties.

In a tolerance space, we can build up a mathematical structure that is similar, in the sense of being able to handle global properties, to general topology. Many of the mathematical structures that can be built up in topological spaces can also be built up in tolerance spaces. Terms and notations used in the mathematics of tolerance spaces were, therefore, borrowed from general topology. For example, we can define precisely tolerance connectivity and tolerance holes. Zeeman (1962) applied algebraic topology, homology theory, to capture global tolerance properties in a discrete set. Given a tolerance space  $(X, \xi)$  we can construct a simplicial complex, consisting of all simplexes, where a simplex is a finite oriented subset of  $X$  all of whose points are within a given tolerance. The homology group  $H(X, \xi)$  refers to the homology group of this complex. Homology theory provides the best way to capture global topological properties. From the homology group, we can capture the global tolerance properties of the tolerance space all at once, that is, from the homology group  $H(X, \xi)$  of a tolerance space  $(X, \xi)$ , we can read off how many tolerance connected components the set  $X$  with tolerance  $\xi$  consists of, and how many holes each of these connected components contains. Zeeman proved that such global topological properties, including tolerance connectivity, the number of tolerance holes, and dimension, are preserved under tolerance homeomorphism. This theorem provides the mathematical basis for defining global properties in a discrete set, and for applying topological invariants to describe perceptual organization.

Let us give a simple example of Figure 17 to illustrate this mathematical description of homology theory. Taking a tolerance of, for example, 1 cm, the homology group of Figure 17 describes the most noticeable global tolerance



**Figure 17.** An example to illustrate the psychological relevance of homology theory of a tolerance space. Taking a tolerance of, for example, 1 cm, the homology group of the dot array describes that the array has two pieces, and one of them has a hole in it. This mathematical description matches our holistic perception of the dot array.

properties: Figure 17 has two pieces (two tolerance connected components), and one of them has a tolerance hole in it. This mathematical description of global properties in the disconnected dot array is consistent with our holistic perceptual organization in this disconnected dot array. This mathematical description matching our holistic perception may seem to be straightforward, but it is by no means trivial. It indicates that the homology theory of tolerance spaces is exactly the mathematics we need to grasp the nature of global properties in perceptual organization. In the following part of this section, I will further clarify one of my major conclusions that holistic perception is essentially the perception of global tolerance properties. That is, Gestalt laws represent an expression of a common underlying principle that perceptual organization is based on global tolerance properties.

### The applicability of tolerance spaces to perceptual organizations

#### *Connectivity differing from spatial connection and global tolerance properties in texture perception*

It is interesting to see that a most interesting and basic concept raised at the very beginning of the study of texture perception is a topological one, namely, “connectivity differing from spatial connection” (Julesz, 1965). Julesz designed some texture demonstrations to examine the following question: Can two textures, even if they are connected in space, be discriminated solely by certain differences in their properties? In one his example, texture discrimination was determined by grouping based on brightness similarity. The random pattern at top left forms two easily discriminated subpatterns, when the subpattern on the left contains mostly black and dark grey dots, similar in brightness to each other, and the subpattern on the right contains light grey and white dots, again similar in brightness to each other. When the dark grey and light grey dots are reversed (top right), so that the dots within each half field are less similar to each other in brightness, the two areas are no longer readily discriminated. Julesz proposed that texture discrimination based on similarity grouping involved a process that may be called “connectivity detection”, which should not be confused with the actual spatial connectivity. He further claimed that this process “connectivity detection” is basic to visual tasks and that it is a more primitive process than form recognition. To conceptualize the process of extracting neighbouring points (that have similar brightness or colours) and forming clusters as the concept of “connectivity detection” can help to provide a deeper explanation of similarity grouping than the “catalogue” approach, characteristic of Gestalt psychology. However, Julesz gave up his effort to provide a precise definition of such “connectivity” (different from the actual spatial connection), perhaps because he proceeded to develop his texton theory (1981), which used a “local-to-global” approach. However, connectivity detection is essentially a global

process. An additional reason may be mathematical in nature. Julesz indeed realized an intrinsic relation between his intuitive “connectivity” based on similarity and the physical connectivity in space. However, he did not find a proper mathematical concept to link them, and, therefore, did not propose his intuitive notion of “connectivity” in relation to topology. I propose that the mathematics of tolerance spaces may provide a way to bridge the gap between Julesz’s intuitive notion of “connectivity” and the mathematical concept of connectivity. As mentioned before, elements of the set of a tolerance space need not to be restricted to physical dots in space, but can also represent states of physical properties (such as brightness and colour), geometric properties (such as length and slope), and other properties that generate similarity groupings. This power of mathematical abstraction gives tolerance spaces the strength to provide a unified description for different forms of grouping, including proximity and similarity based on various factors. To make the discussion less abstract let us take the same texture (Julesz, 1965) as an example to show how the global properties in tolerance spaces can be used to describe “connectivity detection” based on brightness similarity grouping. The elements of the tolerance space for this particular texture discrimination can be considered different brightness values. Thus, the set  $X$  of the tolerance space is the set of brightness values and the tolerance  $x$  can be assigned as all pairs of brightness values less apart than a certain brightness value. The organization based on similarity of brightness can then be described in terms of brightness tolerance connectivity. In the tolerance space of brightness, the dots representing dark grey and black fall within the brightness tolerance, and the dots representing light grey and white also fall within the brightness tolerance. Thus, the dots form tolerance-connected components so that the random pattern was perceived as two connected (easily discriminated) subpatterns. In contrast, the two dots representing black and light grey are far apart and fall beyond the tolerance, and the two dots representing white and dark grey also fall beyond the tolerance. Thus, they are not tolerance connected so that no discriminated subpatterns were perceived. This analysis in terms of tolerance connectivity is applicable to other similarity organizations as well, such as the discrimination of coloured textures, namely random fields of coloured dots (Julesz, 1965). What is needed in this case is to describe the tolerance space in colour space with a colour tolerance.

It is also interesting to see the applicability of tolerance spaces in explaining several seemingly mysterious phenomena in stereoscopic perception with random-dot patterns. Systematic changes of the random dots, such as blurring, size-reducing, and adding noises, were made. Nevertheless these changes did not interfere with the effect of stereoperception. These facts seem not to be predicted by any point-to-point matching. However, the effect is understandable from the perspective of tolerance connectivity: If the changes of blurring, size-reducing, and adding noise all fall within corresponding tolerances, tolerance connectivity of corresponding random-dot field remains. Let us examine one more example

(Julesz, 1965) for which, to my knowledge, there is still no commonly accepted explanation. When the two images of random-dot stereograms are viewed with a stereoscope, a centre panel with a clear-cut contour are seen floating above the background. The interesting finding is that, when the density of the random dots was reduced, within a quite large extent, observers still keep the percept of the clear-cut contour of the floating panel. Only when the density of random dots was reduced even further, observers still had stereoscopic percept of the floating dots but the dots were not integrated into an unified whole anymore. The disappearance of the clear-cut contour reveals that the percept of the clear-cut contour should be regarded as an achievement of the perceptual system under certain conditions. This phenomenon does not follow from any local theory based on point-to-point match. From the viewpoint of tolerance spaces, the explanation is, however, simple and transparent. As soon as the neighbouring dots would fall within the tolerance, the density reduction of the dots would not interfere with the tolerance connectivity of the dots, which produced the subject contour. It is only when the gaps exceed the tolerance, the unified subjective whole made up by the random-dots breaks apart, even though individual dots are still seen floating separately in depth.

In sum, the “connectivity” proposed by Julesz (1965) can be formally defined in terms of tolerance connectivity. Tolerance connectivity shares the same mathematical structure with physical connectivity, which is actually just a special case of the general framework of tolerance connectivity. The global tolerance properties, including tolerance connectivity, serve a theoretical basis for describing perceptual organization in a unified manner.

*Uniform connectedness operates prior to proximity?  
A critical test of the applicability of tolerance spaces to  
perceptual organization*

As discussed in Part I, Palmer and Rock (1994) have proposed a new grouping principle, uniform connectedness (UC), which asserts that a connected region of uniform visual properties strongly tends to be organized as a single perceptual object. They further argued that UC precedes other classical grouping principles, including even proximity. Their claim that UC “can overcome the law of proximity” is particularly critical and actually present a challenge to the topological approach based on tolerance spaces.

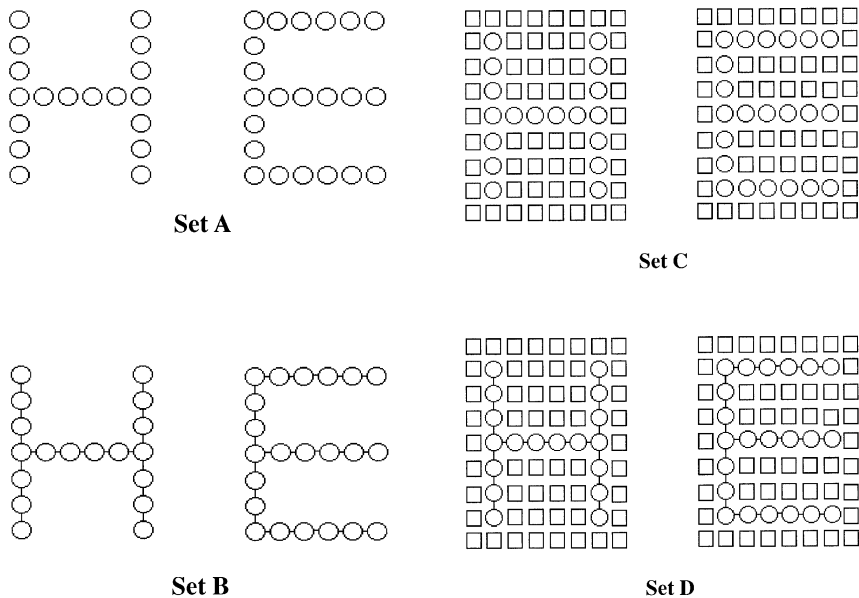
A basic inference from the very nature of the mathematical structure of tolerance spaces is that proximity is, among other factors, the most primary organizational factor. A tolerance is an extremely simple algebraic relation, even simpler or more fundamental than the equivalent relation. Whereas the equivalent relation is defined as reflexive, symmetrical, and transitive, tolerance is only defined as reflexive, and symmetrical but not transitive in general. Thus,

mathematical structures built on tolerance are more primitive than those built on equivalence, such as distance or metrics. In tolerance spaces, no structure of distance or metrics is needed. From mathematical structures of tolerance spaces, proximity requires just that two physical dots are sufficiently close to each other falling within a tolerance, but no distance measurement. Proximity is, therefore, more fundamental than any similarity based on variables required distance or metric measurement. Furthermore, the primacy of proximity also follows the following direct logic. Any similarity grouping based on a feature should occur after detecting or, at least to some extent, processing the feature. For example, the organization based on colour similarity will certainly pre-require at least some processing of the colour. In contrast, proximity requires nothing to process but just physical closeness. In other words, to extract feature values is necessary for any feature similarity, but for proximity no feature (distance) values are pre-required. These analyses, based on the mathematical nature of tolerance spaces and the direct causal relation, lead to a basic constraint on the theory of tolerance spaces, that is, in comparison with various kinds of similarity factors, proximity should play the most fundamental role in organization.

In addition, from the perspective of tolerance spaces, a set of disconnected dots arranged in a row is mathematically equivalent to a solid line-segment, if all their neighbouring dots fall within the same tolerance. That is, the dot row (representing grouping by proximity) and the solid line segment (representing grouping by UC) share the same global property of tolerance connectivity. If perceptual organization is essentially abstracting global tolerance properties, it follows that the dot row and the solid line segment should demonstrate the same grouping efficiency. This prediction may not be consistent with our everyday perceptual experience, but with the topological structure of tolerance spaces. In sum, proximity precedence should be considered as a foundation stone of the theory of topological approach, specifically of applying tolerance (topological) properties to describe perceptual organization. If the stone would be dislodged, the whole “tolerance spaces” building would fall. So, the claim that UC operates prior to proximity, if correct, would be an earthquake to shake the foundation stone.

Fortunately, even though intuitively this claim sounds true, our experimental results have demonstrated that this claim is simply not true; rather, grouping by proximity can be as fast and efficient as that by UC (Han, Humphreys, & Chen, 1999b). For example, the stimulus displays in Figure 18 are used to measure grouping by proximity (set A), grouping by proximity and UC (set B), grouping by similarity (set C), and grouping by both similarity and UC (set D). The subjects were required to discriminate the target letters (H vs. E) regardless of how the small circles were grouped. The results turned out that UC can indeed facilitate performance relative to grouping by similarity. Nevertheless, there was no evidence for stronger grouping by UC than by proximity. Thus, it appears that proximity grouping is at least as efficient as grouping by UC.

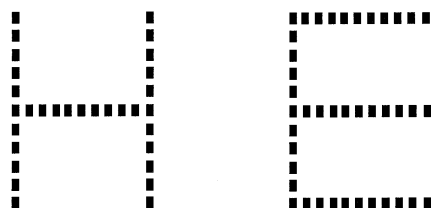




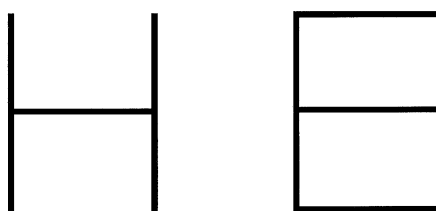
**Figure 18.** Compound stimuli used to compare grouping effects based on proximity, similarity, and/or UC. The small circles were grouped by proximity (Set A), both proximity and UC (Set B), the similarity based on the shape of circle in comparison with the shape of square (Set C), and both the similarity and UC (Set D).

One more experiment testing this claim was conducted. As shown in Figure 19, two sets of large letters (H and E) were made up of solid lines or separated small solid rectangles. This allowed a test to be made between elements grouped by proximity and a solid-line baseline condition, to give the best chance for precedence of grouping by UC over that by proximity. To prevent subjects perceiving the gap between two adjacent rectangles, the stimuli were presented briefly and masked; a control condition was also administered in which subjects had to discriminate whether the target letters were made up of solid rectangles or were merely solid lines. The experiment again did not show any difference between the stimuli formed by UC and by proximity, contrary to the assertion that UC operates prior to proximity. These experimental results appear to be counterintuitive. Nevertheless, from the mathematical structure of tolerance spaces, UC is just a special case of proximity. It is, therefore, not a surprise that grouping by proximity can be as fast and efficient as that by UC.

The primacy of proximity has also been supported by a systematical empirical study (Kubovy, Helcombe, & Wagemans, 1998). Kubovy and co-workers measured proximity-grouping strength and the effect of configuration in their experiments, in which they briefly presented multistable dot patterns that can be perceptually organized into alternative collections of parallel strips of dots, and



Set A



Set B

**Figure 19.** Stimuli used to compare grouping by proximity with that by UC. The two sets of large letters (H and E) were made up of solid lines, as a baseline (Set B) or separated small solid rectangles (Set A).

parametrically varied the distances between dots and the angles between alternative organizations. They found that surprisingly, the configural properties that were systematically varied did not affect grouping by proximity, and the relative strength of proximity grouping approximates a decreasing exponential function of distance between dots, which was also robust under transformations of scale in space and time. In addition to that these data strongly support the primacy of grouping by proximity, it is worth emphasizing that the negatively, and rapidly decreasing function (rather than a linear one) of distance between dots matches well with the concept of distance tolerance, which implies an on–off (rather than a gradually changing) range within which distance changes are ignored.

#### PART IV: TOPOLOGICAL STRUCTURE AND FUNCTIONAL HIERARCHY IN FORM PERCEPTION—A NEW ANALYSIS OF PRIMITIVES OF VISUAL PERCEPTION

To complete the theory of topological perception, one needs to address the relationship between the perception of topological and other geometric prop-

erties. It is obvious that, even if topological perception is fundamental to vision, it cannot be the whole story. Otherwise, we would not be able to tell a teacup from a doughnut! I now proceed to discuss experiments on how topological perception relates to perception of other geometrical properties. These experiments reveal a functional hierarchy, in which the relative salience or priority in the perception of different geometric properties is remarkably consistent with the hierarchy of geometries, stratified by Klein's Erlangen Program. As a consequence, I first introduce the idea of Klein's Erlangen Program, and outline its relevance to the study of the primitives of visual perception.

### **Klein Erlangen Program: A mathematical framework for describing the relations between topological and other geometrical properties with respect to form stability**

One important factor in evaluating the potential primitives for perceptual representation is their relative stability. When an object in the natural environment, say a flying bird, is subjected to changes due to its motion (which may be nonrigid), or to changes of illumination, it is generally the case that some of its form properties may be altered, while others (e.g., its identity) remain invariant. Is there any formal way to clarify the relationship between the various invariants with respect to their levels of structural stability?

The mathematician Klein, in 1872, proposed a general principle for constructing different geometries that is now known as the Erlangen Program, in which a geometrical property is considered as an invariant over a corresponding transformation group. For understanding the relations between different geometrical properties, the most useful idea rooted in the Program is that the more general a transformation group, the more fundamental and stable the geometric invariants over this transformation group. Using this principle, Klein built a hierarchy of geometries, stratified in ascending order of stability: Euclidean geometry, affine geometry, projective geometry, and finally topology with the highest stability. Within the hierarchy of geometries, topological properties are the most stable, because the topological transformation group, "one-to-one and continuous" transformations, among others, is the most general. Alternative geometries can be devised for which constraints on a corresponding transformation group are stricter. The projective transformation group is less general (it is a linear transformation group) than the topological transformation group but more general than the affine transformation group. Therefore, projective invariants are less stable than topological invariants, but relatively more stable than affine invariants. Affine geometry is formulated by adding more constraints (e.g., that makes the parallelism invariant) to the projective transformation group so that affine invariants are less stable than projective invariants but relatively

more stable than Euclidean geometry, which is devised by adding more constraints (e.g., that make length invariant) to the affine transformation group. Thus, the Klein Program provides a formally precise way to stratify the relations between different geometrical invariants with respect to their relative structural stability. Table 3 illustrates different invariants in the hierarchy of geometries, based on the contrasting transformation groups, in an ascending order of relative form stability: Typically, from point position to mirror symmetry, to line orientation, to collinearity, and to connectivity and holes.

In the following sections, experimental results collected from various topics in the study of perceptual organization will demonstrate the psychological reality of Klein's mathematical structure, and reveal a functional hierarchy of form perception that is remarkably consistent with this stratification of geometries.

### Configural superiority effects and a systematic investigation of the relations between topological perception and perception of other geometric properties

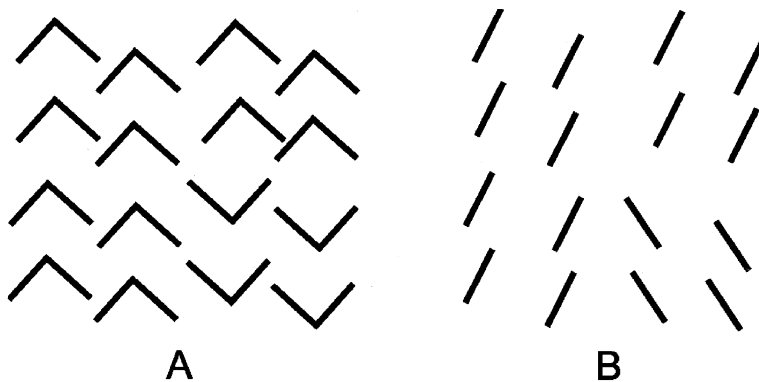
The feature-analytic theory of visual perception always posits certain simple, local parts as the primitives of figure perception. Perhaps because straight line segments are readily accepted as one of the simplest component parts of figures, it has often been taken for granted that line segments provide basic stimuli in experimental studies of visual form perception. A large set of data, particularly the neurophysiological findings of detectors of lines and edges in the mammalian cortex by Hubel and Wiesel (e.g., 1965, 1968), seem to provide strong support for line-based shape coding. As is well-known, in Marr's computational model of vision (1982), Hubel and Wiesel's finding of line-segment detectors was considered as the physiological basis of "the primal sketch", which describes the early processing of vision as the extraction of the intensity changes represented, typically, by line segments with their local geometry.

TABLE 3  
Invariants of geometrical groups

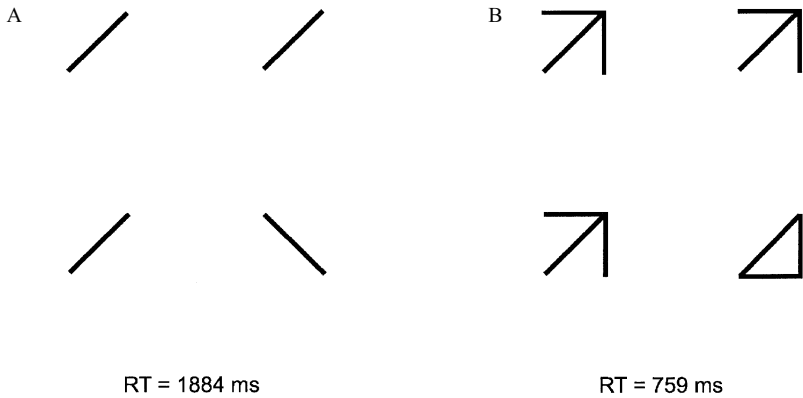
<i>Geometrical groups</i>	<i>Invariant properties</i>
Identity group	Point position
Mirror-symmetry group	Mirror-symmetry
Displacement group	Line orientation
Congruence group	Line length
Similarity group	Parallelism
Projective group	Collinearity
Topological group	Connectivity and holes

Behavioural studies, along with neurophysiological studies, also made efforts to search for psychological evidence supporting the line-based shape coding. Beck and his associates (1967, 1972; Beck & Ambler, 1973), and others measured the discriminability of various featural factors. An example of their results shown in Figure 20 (adapted from Olson & Attneave, 1970) represents one of their major findings. In Figure 20B, the disparate quadrant differs from the rest in line slope, while in Figure 20A, in contrast, the disparate quadrant contains exactly the same line segments as the rest but differs in direction pointed by the right angles. The result, that the odd quadrant was detected much faster in Figure 20B than in Figure 20A, was taken as support for line-based models of shape coding.

However, such a feature-analytic model has been challenged by findings of contextual effects in form perception, for example, the object superiority effect (Weisstein & Harris, 1974) and configural superiority effects (Pomerantz, Sager, & Stoeve, 1977). A typical example of configural superiority effects is illustrated in Figure 21. While in Figure 21A the line segment in the odd quadrant differs from the rest, in Figure 21B all four quadrants each contain exactly the same three line segments. If it was true that visual processing began with an analysis of a figure into line segments, the discrimination in Figure 21A would be easy, being based at a primitive stage; with the stimuli in Figure 21B, discrimination should be more difficult since the same primitive line detection will be activated in all cases. However, the result turned out to be just the opposite to this line-based prediction, with the stimuli in Figure 21B being much easier than those in Figure 21A. This “configural superiority effect” indicates that configural relations between line segments rather than line segments themselves



**Figure 20.** Stimuli used as evidence for line-based models of shape coding (adapted from Olson & Attneave, 1970). In (B), the disparate quadrant differs from the rest in line slope, while in (A), in contrast, the disparate quadrant contains exactly the same line segments as the rest but differs in direction pointed by the right angles.



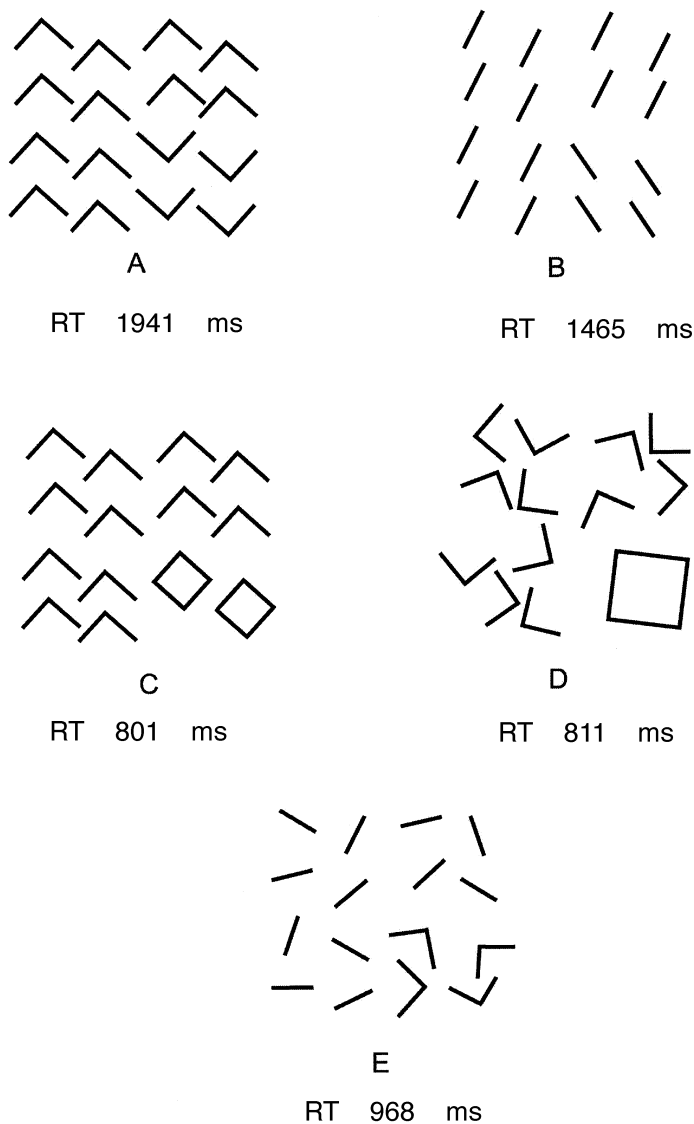
**Figure 21.** A counterexample against the line-based model of shape coding. In (A), the disparate quadrant differs from the rest in line slope, while in (B), in contrast, the triangle in the disparate quadrant contains exactly the same line segments as the arrows in the rest. Also shown are the mean RTs for each stimulus display.

may play a basic role in visual processing, and thus provides a major challenge to the line-based model of shape coding.

The feature-analytic theory makes no provisions for the nature of contextual effects, and thus is unlikely to be able to explain the powerful influence of contexts in visual processing. Pomerantz (1978) made an important claim that “The sloped line detectors, of the variety discovered in the mammalian cortex by Hubel and Wiesel, are not the primitives of the human visual pattern recognition system.” However, the issue remains of exactly what kinds of form information produce the superiority effects? This question, concerning configural superiority effects, was raised nearly 30 years ago but is almost completely ignored now. Nevertheless it raises the fundamental question of “Where to begin” in visual processing.

From the point of view of topological perception, the superiority effect demonstrated by the stimulus of the “triangle–arrow pair” simply demonstrates the superiority effect for perception of holes over individual sloped lines. Using configural superiority effects as an index, a testable framework can be advanced to address the fundamental question: What are the primitives of visual perception? As shown in Figure 22, five stimulus arrays were designed to measure the relative salience of different geometric properties in a texture-segregation task (Chen, 1986).

- *Stimulus A* represents texture discrimination based on a difference in orientation of angles, a kind of Euclidean property. But its each quadrant contains the exactly the same line segments.



**Figure 22.** The five stimulus arrays designed to measure, in a texture-segregation task, the relative salience of the different levels of geometrical invariants stratified by Klein's Program. They represent discriminations based on (A) a difference in orientation of angles, a kind of Euclidean property, (B) a difference in parallelism, a kind of affine property, (C) the topological difference in holes (in the odd quadrant, two of the four right angles form two squares), (D) also the topological difference in holes (in the odd quadrant four right angles form a larger square), and (E) a difference in collinearity, a kind of projective property. Also shown are the mean RTs for each stimulus display. (See text for details.)

- *Stimulus B* represents discrimination based on a difference in parallelism, a kind of affine property. The line segments in the odd quadrant are parallel but differ in parallelism from the rest.
- *Stimulus C* was adapted from stimulus A: In its odd quadrant, two of the four right angles are rotated  $180^\circ$  so that they could join another two right angles to form two squares, while the four right angles in each of the rest quadrants are open.
- *Stimulus D* was designed so that the odd quadrant contains a larger square formed by four right angles, but the rest each contain four right angles distributed in random orientations. Thus, stimuli C and D both were designed to represent discrimination based on holes.
- *Stimulus E* represents discrimination task based on a difference in collinearity, a kind of projective property. Its odd quadrant contains bent lines, while in the remaining quadrants, there are straight lines.

Subjects were asked to report which quadrant was different from the rest. The results, shown also in Figure 22, demonstrate a clear-cut correlation between RTs and the different levels of geometrical invariants stratified by Klein's Program. First, RTs with C and D, both representing discrimination based on holes, were significantly shorter than those with other stimuli, but the difference in RTs between C and D did not reach significance. Second, there was a collinearity superiority effect (E) over parallelism and orientation (B and A). Third, discrimination based on parallelism (B), was superior to that based on orientation (A).

Let us consider the various attributes that could be proposed to explain these data.

*The closure-superiority effect.* Since the odd quadrant in Figure 22C contains the same sloped lines and the same right angles as the rest, neither the line-based nor the angle-based feature-analytic model can explain the performance. However, the odd subarray in Figure 22C is topologically different from the rest in holes. This gives a direct explanation for the data. With respect to Figure 22D, since each quadrant in stimulus D contains the same number (four) of vertices and the same right angles, and there is no meaningful difference in the orientations of the angles in the four quadrants. Hence neither the line-based, the angle-based, nor the vertex number-based feature-analytic models can explain this superiority effect. Nevertheless, it is interesting to note that although, from the perspective of a line-based model, D is quite different from C, the difference in RTs between them did not reach significance.

*The relative superiority effects of collinearity over parallelism.* In Figure 22E there is no meaningful difference in line slopes between the odd quadrant and the rest. Hence if the discriminability could be explained by line slope, E



would be less discriminable than B, or at least the same as A. The reverse result we observed is consistent with the stratification of geometries: Projective geometry (collinearity) is relatively more basic than affine geometry (parallelism).

*The relative superiority effects of parallelism over orientation.* Figure 22B was previously used to assess discrimination based on line slope. The idea of invariants over transformations, however, suggests a new analysis of the data. The fact that line segments share the same slope essentially implies that they are parallel, and the discrimination in Figure 22B, therefore, may be actually based on parallelism. Thus, the faster discrimination of B than that of A may be explained in terms of relative superiority of parallelism over orientation of angles—a Euclidean property.

These data obtained with the five stimuli demonstrate a complete match between the order of reaction times and the stratification of geometries with respect to their relative form stability. In terms of the fundamental question of “what are the primitives of form perception”, the above discussions consistently lead to the general framework proposed at the beginning of the present paper, namely, that a primitive function of the visual system is the perception of global topological properties, and that with respect to the time dependence of perceiving different geometrical properties, there is a functional hierarchy of form perception similar to the gradation of geometry stratified by Klein’s Erlangen Program. The primitives of form representation are considered as invariants at different geometrical levels, including the basic level of topology. In the following sections, I will discuss other pieces of evidence supporting the functional hierarchy.

### The global precedence hypothesis and the functional hierarchy of perceptual organization

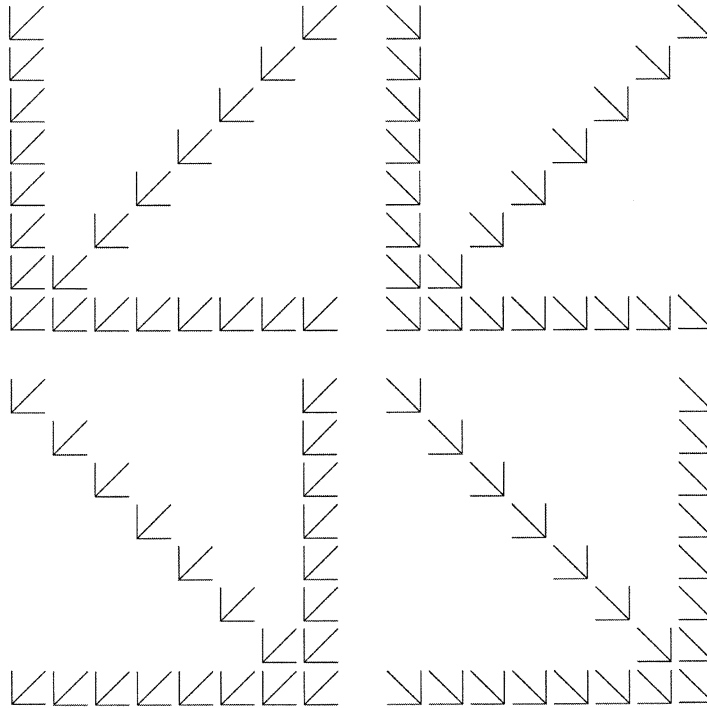
Navon (1977) reported that subjects responded faster to a compound letter relative to component letters, and that the compound letter interfered with responses to component letters when the two levels were incompatible, but not vice versa. On the basis of the performance advantage for compound letters, Navon put forward a “global precedence hypothesis”, popularized by the metaphor of “forest before trees”. However, more recent research reveals that there is considerable variability in both “global” RT advantage and the “global-to-local” interference effect. For example, each effect varies as a function of the absolute size of the compound letters (e.g., Kinchla & Wolfe, 1979), the density of the component letters (e.g., LaGasse, 1993), and spatial frequency components contained in the compound stimuli (e.g., Hughes, Norzawa, & Kitterle, 1996; Lamb & Yund, 1996).

According to the proposed functional hierarchy of perceptual organization based on topology, the performance advantage for compound letters is quite straightforward: Global precedence reflects the primacy of proximity in perceptual organization. As can be seen in the stimulus displays used by Navon (1977) and others, the compound letters were formed, probably, by proximity. Early grouping based on proximity would generate both the “global” (more precisely “compound”, as discussed in Part V) RT advantage and the “global-to-local” (more precisely, compound to component) interference. The effect of other variables (such as absolute size, density, exposure duration, and spatial frequency component) arise because these variables affect proximity grouping (e.g., proximity effect decreased when letter density was lowered). However, in testing this explanation we face the difficulty that proximity is often confounded with similarity. In order to prove the major role of proximity in the “global precedence effect”, we must find a way to dissociate proximity from similarity.

A technique for dissociating proximity from similarity was devised in the following experiments (Han, Humphreys, & Chen, 1999a). As shown in Figure 23 an arrow-like figure was used, instead of letters, to serve a component pattern for the compound stimuli. The directions of the compound arrows were either consistent or inconsistent with that of the component arrows. Subjects were asked to discriminate the direction of an arrow at the compound or component levels. Like the compound letters used before, the compound figures also demonstrated the “global precedence effect” and the “global” interference without “local” inference (Han, Humphreys, & Chen, 1999a).

In order to eliminate the effect of proximity, the compound arrows were put against a background of cross elements, as shown in Figure 24. The distance between an individual component arrow and its surrounding crosses was equal to that between two adjacent component arrows. Under this circumstance, the coding of the compound arrow shape based on proximity should have been reduced (i.e., the component elements making up the compound arrow should have been no closer to each other than they were to the surrounding background crosses), and the dominant grouping factor defining the stimulus became similarity of shape instead of proximity. Thus, if the explanation of the “global precedence effect” based on priority of proximity is valid, we would expect that the compound figures of Figure 24 would not show “global precedence” any more. This prediction was verified. The global advantage observed with Figure 23 was eliminated: “Global” RTs were no faster than “local” RTs, and “local” interference was stronger than global interference (Han, Humphreys, & Chen, 1999a).

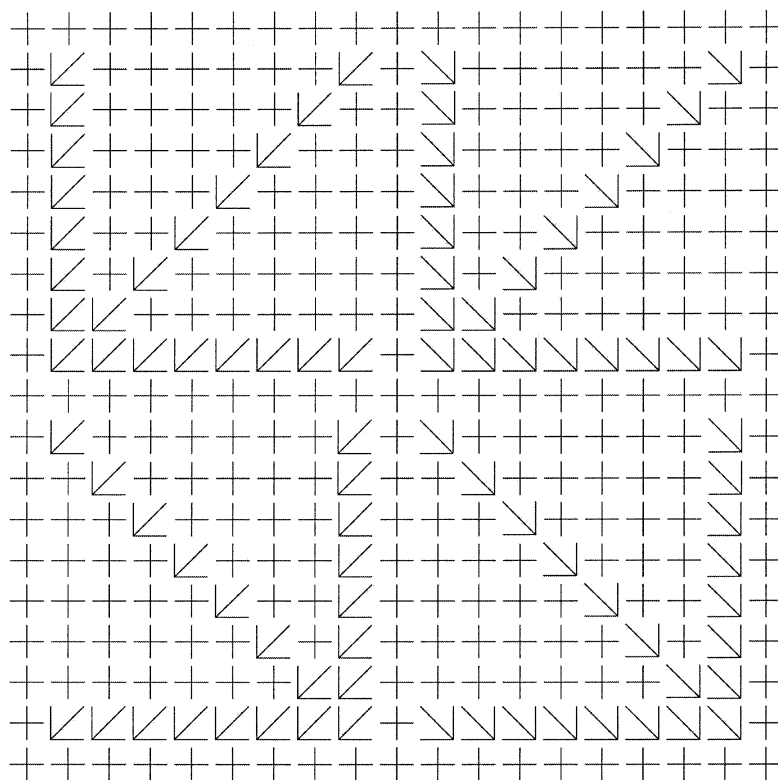
A further control experiment was designed to test for contrasts between grouping based on topological similarity and on similarity of local geometrical properties. In Figure 25, the compound figures were the same as those of Figure 24 except that square elements were used to form the background instead of the cross elements. In this case, there was a topological difference between the open



**Figure 23.** The compound arrows, in which a small arrow-like figure was used to serve a component pattern. Discriminations of the direction of an arrow were made at the compound or component levels.

arrow figures and the closed square elements. In terms of the proposed functional hierarchy of perceptual organization, we would expect that the compound figures formed by topological similarity (Figure 25) to show some precedence in comparison with the compound figures formed by local geometrical similarities (Figure 24). The results showed that grouping based on difference in holes (between the component arrows and the background squares) facilitated the perception of compound figure, resulting in a reduction of the local advantage (Han, Humphreys, & Chen, 1999a).

The above results indicate that the performance advantage for the compound forms respects the principle of the early topological perception and the functional hierarchy of grouping. In addition, these results also raise interesting issues about the relation between concepts of the “global” and the “compound” (or “whole”), which are used often interchangeably in the literature on “global precedence effects”. The paradigm using compound letters provides an operational technique to examine the relations between proximity and

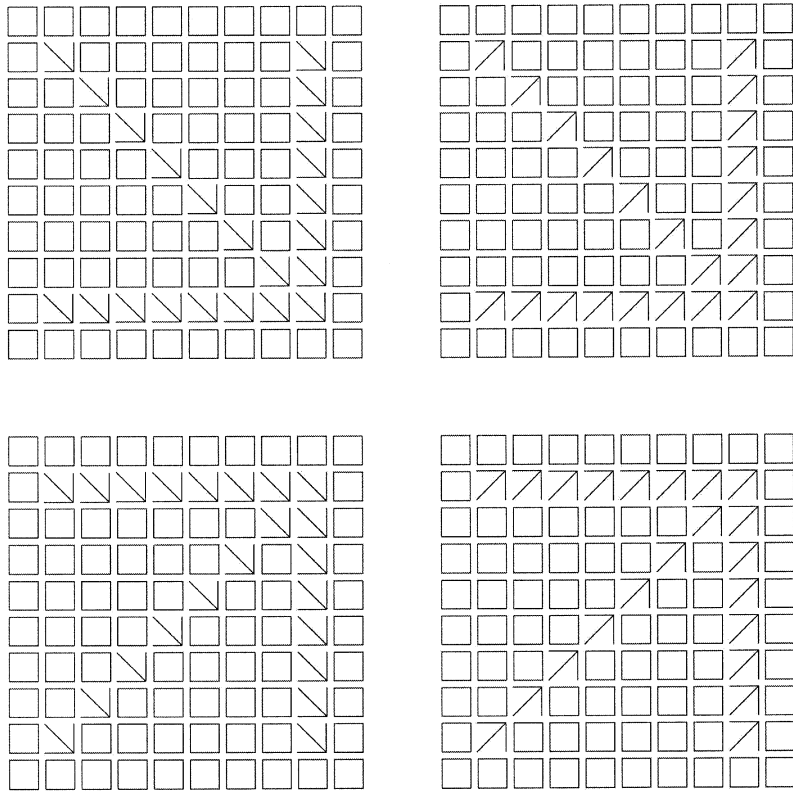


**Figure 24.** The technique for dissociating proximity from similarity. In order to eliminate the effect of proximity, the compound stimuli are made by embedding the stimuli of Figure 23 in a background of small crosses.

similarities, and the claim of “global precedence” touches the marrow of the Gestalt tradition. However, our present data suggest that the claim of “global precedence effects” is valid when one understand the “global” in terms of abstract topological invariants, and we cannot simply claim “compound (or whole) precedence effects”. These issues are related to a basic issue in the study of perceptual organization: How to define the concepts of “global vs. local”, and will be further discussed in the section of Part V regarding Gestalt psychology.

### The relative salience of geometric properties for 3-D form discrimination

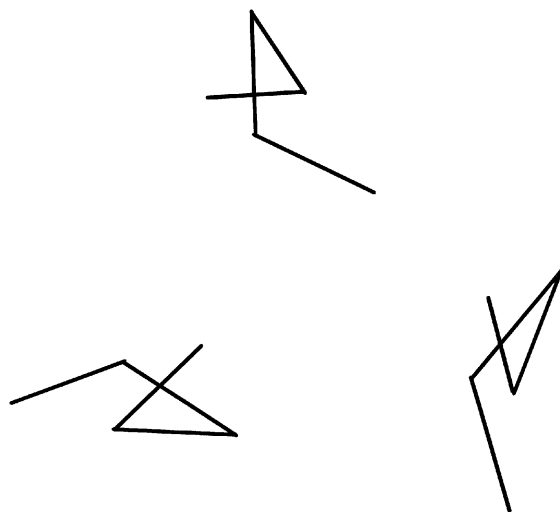
So far the stimulus figures used in the above experiments were all 2-D forms. Todd et al. (1998) extended the study of the functional hierarchy from 2-D to 3-



**Figure 25.** The compound stimuli designed to further test for contrasts between groupings based on topological similarity and on similarity of local geometrical properties. In the case that the square elements were used to form the background instead of the cross elements, the compound figures are formed by grouping based on topological difference in holes (between the component arrows and the background squares) rather than the local feature of arrow directions.

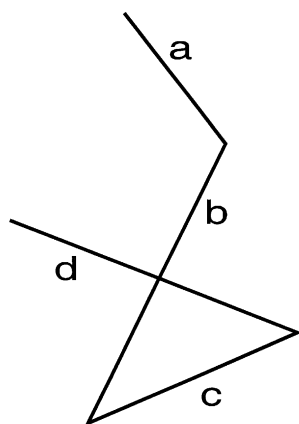
D forms. This study examined the importance of invariance under change for the binocular perception of configurations of line segments in 3-D space. The displays were viewed through LCD shuttered glasses that were synchronized with the monitor's refresh rate.

Each stimulus display contained a triangular arrangement of three wire-frame objects (Figure 26), all of which contained four connected line segments. The upper object in the triangular configuration was designated as the standard, and the two lower objects were designated as test figures. The projected images of these objects, on the plane of the display screen, all had the same 2-D topology as shown in Figure 27. The four line segments—labelled in the figure as a, b, c,



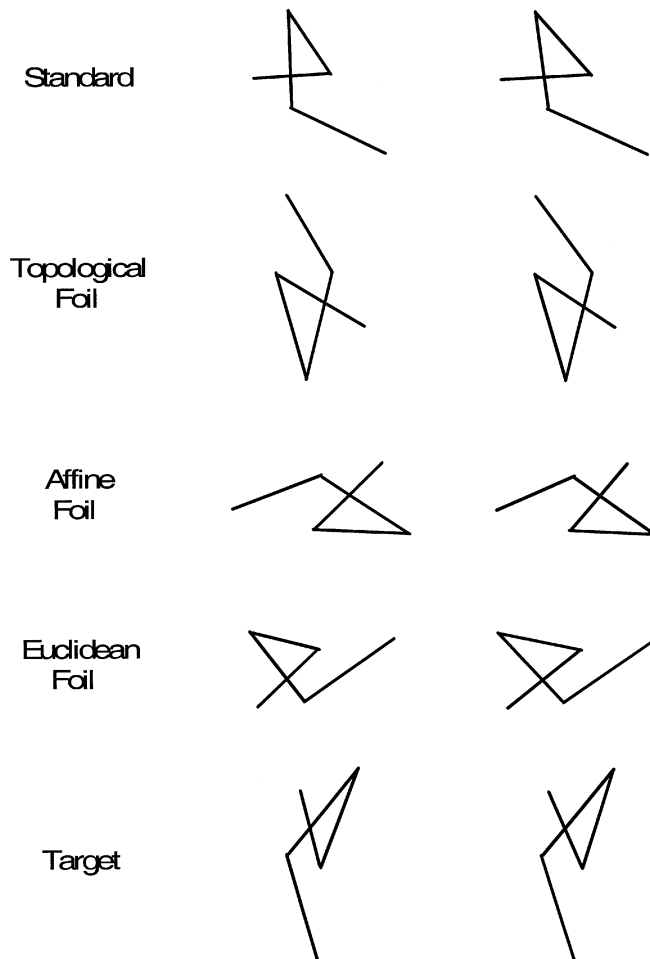
**Figure 26.** An example stimulus configuration with a standard (upper) and two possible test wire-frame objects (lower). (See text for details.)

and d—were oriented in three-dimensional space so that b was connected to a and c; c was connected to b and d; and the projected images of b and d intersected one another. In addition, the angle between adjacent segments and between each segment and the line of sight was always greater than  $25^\circ$ . Within those constraints, the lengths and angles of the standard object were generated at random on each trial.



**Figure 27.** Each standard and test wire-frame objects contained four connected line segments (a, b, c, and d) and their projected images all had the same 2-D topology.

There were three distinct experimental conditions, in which the topological (the patterns of co-intersection), affine (coplanarity), and Euclidean relations (the relative orientations of line segments) between the targets and the foils were manipulated. In the *topological condition*, the foil was created from the standard by bending line segment d, as shown in Figure 28, so that the two test figures had different 3-D topologies. On half the trials, the standard was constructed such that d intersected b in three-dimensional space, and the foil was structured such that d was slanted by  $40^\circ$  relative to the plane defined by b and c. On the



**Figure 28.** An example stereogram of a possible standard wire-frame object, a target, and three possible foils from the topological, affine, and Euclidean conditions.

remaining trials this relationship was reversed. That is to say, *b* intersected *d* in the foil but not in the standard. In the *affine condition*, the foil was created from the standard by bending line segment *a* so that the two test figures differed in the affine property of planarity. Segments *b*, *c*, and *d* were always coplanar in this condition. On half the trials, the standard was constructed such that *a* was coplanar with the other segments, and the foil was structured such that *a* was slanted by  $40^\circ$  relative to the plane defined by *b*, *c*, and *d*. On the remaining trials this relationship was reversed—i.e., all of the segments were coplanar in the foil but not in the standard. Finally, in the *Euclidean condition*, the foil was again created from the standard by bending line segment *a*. In this case, however, there were no constraints on planarity or on whether *b* intersected *d* in three-dimensional space. Some example objects from these different conditions are shown in Figure 28. The upper and lower stereograms show a possible standard object and target at different 3-D orientations. The middle three stereograms depict potential foils for this object in the topological, affine, and Euclidean conditions. Note that it is virtually impossible to distinguish these objects from their 2-D projections in the image plane, but that they are relatively easy to discriminate when viewed stereoscopically.

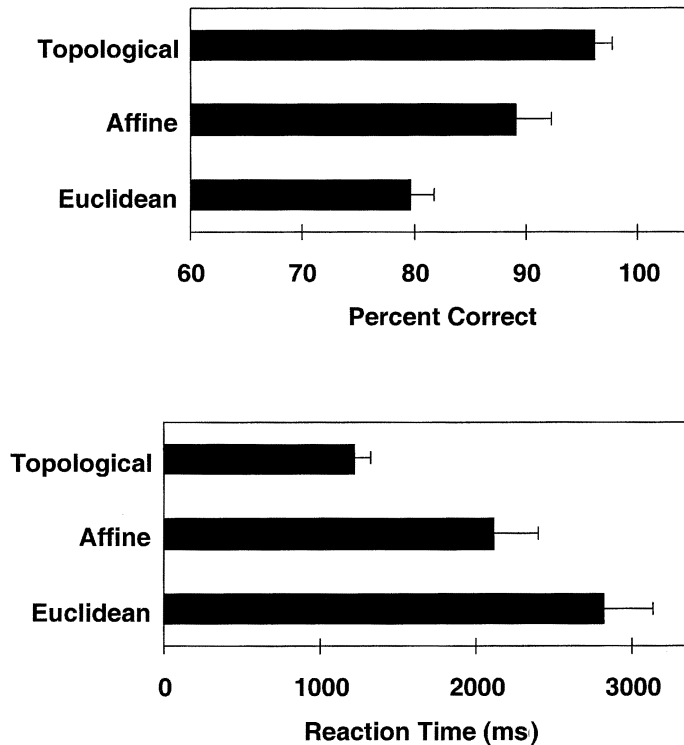
Subjects were instructed to indicate which of the test figures had the same 3-D structure as the standard by pressing the left or right button on a hand-held mouse, and to make their judgements as quickly and accurately as possible. The accuracy and reaction time of each response were recorded, and Figure 29 shows the percentage of correct responses as well as reaction times in the Euclidean, affine, and topological conditions averaged over all 10 observers. Performance was significantly different across the three conditions, and that all pairwise comparisons were significant as well. As is evident from the figures, both accuracy and reaction times provide a clear indication of relative difficulty in the three conditions, the structural discrimination being easiest in the topological condition.

### Neural correlate of form invariants in long-range apparent motion: fMRI evidence for the functional hierarchy

Recently we used fMRI to measure cortical activation to long-range apparent motion, and found that long-range apparent motion activated the anterior-temporal lobe in the visual ventral pathway, and the response varied according to the form stability, in a manner similar to Klein's Program (Zhuo et al., 2003).

As discussed before, the phenomenon of shape-changing transformations observed in apparent motion raises a key question in understanding of apparent motion, namely which invariants of an object are preserved under shape-changing transformations and are consequently used to achieve a correspondence match by the visual system (Chen, 1985). No one has yet devised a general-





**Figure 29.** (1) The percentage of correct responses, and (2) the mean reaction time for correct responses, both averaged over 10 observers in the topological, affine, and Euclidean conditions.

purpose theory of long-range apparent motion that can account for its ecological functions (Palmer, 1999). The current dominant view considers apparent motion to be detected by the same visual channels as real motion (Anstis, 1986), with form vision playing at best a minor role. In particular, it is thought that structural forms are not the correspondence tokens in apparent motion (Ullman, 1979). Nevertheless, this analysis of the correspondence problem in shape-changing, long-range apparent motion and the behavioural finding of topological preference in long-range apparent motion (Chen, 1985), discussed in Part II, led us to hypothesize that long-range apparent motion is actually associated with global form perception.

To test this idea, we used fMRI to investigate human cortical areas mediating long-range apparent motion. The activation stimulus was two squares separated by about  $10^\circ$  and presented in alternation so as to produce apparent motion. The baseline stimulus was the same two squares but presented simultaneously, so that flicker but no apparent motion was perceived. The data revealed activation

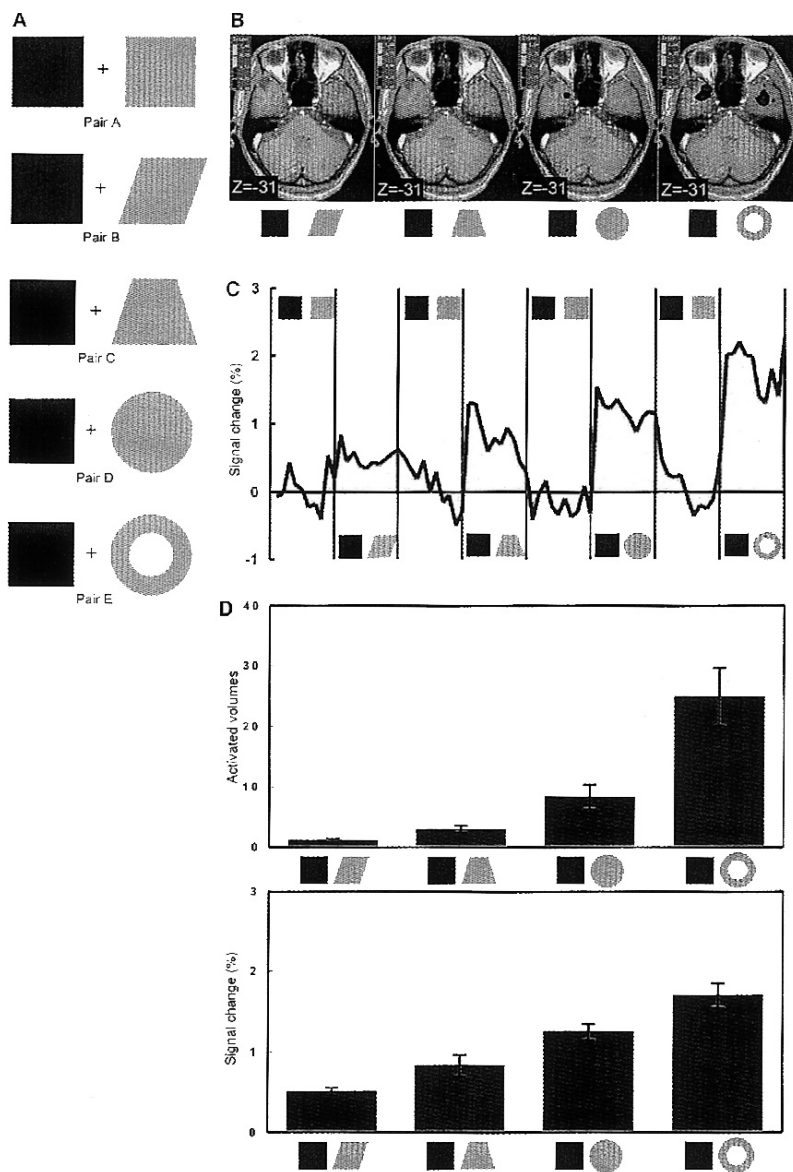
not only in lateral occipitotemporal cortex, but also in the anterior temporal gyri. This latter result was a surprise because the anterior temporal lobe is a late destination of the visual form pathway (Ungerleider & Minski, 1982), which is anatomically far removed from area MT and the where pathway.

If this activation of the anterior temporal lobe indeed occurred because long-range apparent motion is associated with form perception, then varying the form properties should influence the activation pattern. Five pairs of figures (including the baseline pair A) were designed to produce apparent motion (Figure 30A). The differences between the two figures in pairs B to E represent different levels of form stability. In ascending order from pairs B to E, they differ in Euclidean geometry, affine geometry, projective geometry, and finally topology with the highest stability. These constitute a hierarchy of geometries according to Klein's Erlangen Program.

The fMRI data shows that the activated cortical volumes as well as the amplitudes of signal changes in the anterior temporal lobe increased monotonically with increasing levels of the stability of structural differences in the forms (Figure 30B–D). This result suggests that as the stability of structural difference between two forms is increased, the greater the magnitude of cortical activation required to produce the perception of apparent motion between the two forms. Specifically, pairwise comparisons by means of GLM analysis indicated that pair E, representing the highest stability (topological difference), caused the strongest activation in comparison with pairs D, C, and B (for activated volumes,  $p < .016$ ,  $.006$ , and  $.004$ , respectively; for signal intensities,  $p < .012$ ,  $.001$ , and  $.001$ , respectively); pair D, representing higher stability than pairs C and B, caused greater activation than pairs C and B (for activated volumes,  $p < .038$  and  $.019$ , respectively; for signal intensities,  $p < .006$  and  $.001$ , respectively); and pair C, representing relatively higher stability than pair B, caused relatively greater activation than pair B (for activated volumes,  $p < .020$ ; for signal intensities,  $p < .030$ ).

Two additional experiments further tested the neural correlate of form stability. In one experiment, in producing long-range apparent motion, two pairs of figures were compared with the baseline task of two identical S-like figures (Chen, 1990). The S-like figure is topologically different from the ring in holes. However, the S-like figure was made to approximate the area of the ring. Even though the local-feature differences between the S-like figure and the ring (such as luminous flux difference, spatial frequency components, and perimeter length) were minimized, this topologically different pair caused stronger activation in the anterior temporal lobe than the topologically equivalent pair of S-like figure and disk (for activated volumes,  $p < .03$ ; for signal intensities,  $p < .01$ ).

Kinetic shapes, instead of luminance-based ones, were used in another experiment. Kinetic shapes are defined solely by spatiotemporal correlations rather than luminance differences. Thus all luminance-based features were well



**Figure 30.** (A) The baseline Pair A, and four activation stimulus Pairs B–E that were used to produce apparent motion. Pairs B–E represent different levels of form stability: In ascending order from Pairs B–E, they differ in Euclidean geometry, affine geometry, projective geometry, and finally topology with the highest stability. (B) Activation patterns responding to each of these activation stimulus pairs. (C) The time course of signal changes. (D) The histograms of activation volumes and amplitude of signal change, averaged over the 11 subjects in the anterior temporal lobe for each of these activation stimulus pairs. (See text for details.)

controlled. Recent fMRI studies found that dynamic random dots with 0% coherence produced weak activation in human MT+ (Rees, Friston, & Koch, 2000), which makes the kinetic shapes a particular good choice to control for any confounding effects of differences in total motion energy. Furthermore, the similarity of the apparent motion percepts generated by kinetic forms and luminance-contrast forms affords a test for the generality of this neural correlate of structural stability. In this experiment, a kinetic disk paired with a kinetic triangle, and a kinetic disk paired with a kinetic ring, were compared with two identical kinetic disks, all producing long-range apparent motion. The anterior temporal lobe was activated more strongly by the kinetic forms differing in holes (for activated volumes,  $p < .04$ ; for signal intensities,  $p < .01$ ), similar to what we found with luminance-defined patterns.

Regarding the function of long-range AM, the neuroimaging finding supports the idea suggested by behavioural findings of topological preference in apparent motion reported about twenty years ago (Chen, 1985), that long-range apparent motion may work by abstracting invariants of form, which can be defined by their form stabilities stratified by Klein's hierarchy of geometries.

In summary, these data collected within a variety of approaches and paradigms in the study of perceptual organization reviewed in Parts II–IV consistently support the topological structure and functional hierarchy in form perception, proposed at the beginning of Part I. This novel framework, in which the primitives of visual form perception are considered to be geometric invariants at different levels of structural stability (as opposed to simple components of objects, such as line-segments), highlights holistic transformations and invariance perception rather than discrete feature analysis (which is so popular in current study of visual perception). It has not been escaped my notice that this consequentially raises interesting issues related to major theories of visual perception—some of which will be discussed next.

## PART V: RELATIONS TO MAJOR THEORIES OF VISION

### Gestalt psychology

So far, I have outlined the topological approach to perceptual organization, and showed how it may serve as a formal and testable framework for Gestalt principles. Various critical issues in Gestalt psychology, such as grouping, may be consistently described in terms of extracting global topological (tolerance) invariants in a discrete set. Furthermore, the relationship between various organizational factors may be examined from the perspective of the functional hierarchy, and summarized in the claim that proximity is prior to similarity, and topological similarity is prior to similarity based on other local geometric properties.

The topological approach provides a new analysis of the part-whole relationships, the central concern of Gestalt psychology. Historically, the attitude of Gestaltists towards the concept of “whole” is difficult to grasp. In the psychological literature, part-whole relations are also linked to contrasts between “large-scale vs. small-scale”; “holistic vs. analytic”; “undifferentiated vs. differentiated”; “compounds vs. components”; “integral vs. separable”; “integral vs. nonintegral”; “nonanalysable vs. analysable”. This variety of terms indicates the lack of a clear understanding what “wholes” and “parts” are. To define scientifically the concepts of “wholes” and “parts” remains a major challenge in establishing a formal description of perceptual organization.

### *Global and local*

The topological approach may provide a proper treatment on part-whole relationships: Using the notion of “global and local” in place of “wholes and parts”. From the perspective of transformation and invariance perception under transformation, the term “global” used here refers to being independent of detailed changes. In other words, “global” is linked to invariance or stability preserved under transformations.

A property is considered more global the more general the transformation group is, under which this property remains invariant. Relative to geometrical transformations, topological transformations are the most general and hence topological properties are most global.

On the other hand, from this perspective, “global” and “local” are relative concepts. For example, in comparison with topological properties, projective properties are local, because the topological transformation group is more general than the projective transformation group; but in comparison with affine properties, projective properties are global, because the projective transformation group is relatively more general than the affine transformation group. That is, whereas projective properties are less robust or less stable in resistance to changes than topological properties, they are relatively more robust or more stable in their resistance to changes than affine properties. In the same sense, whereas affine properties are local in comparison to projective properties, they are global in comparison to Euclidean properties. This implication of the definition of “global vs. local” can be grasped by comparing this concept with other concepts, such as “large-scale vs. small-scale”, and “wholes vs. parts”.

**“Global and local” vs. “large-scale and small-scale”.** One pair of concepts that is often confused with “global and local” is “large-scale and small-scale”, which often appear in the so called multiscaled or coarse-to-fine analyses in computer science and artificial intelligence. While spatial scale is intuitively important in visual perception, it is certainly not valid to claim “large-scale precedence”, because no guarantee could be made that a large-

scaled object would be perceived or processed prior to a small-scaled one, or that the perception of a small-scaled object would depend on the perception of a large-object, even though the small one would be a part of the large one. From the perspective of “global vs. local”, it turns out that a large-scaled object possesses both global properties and local properties, whereas a small-scaled object possesses its local properties as well as its global properties. For example, the famous quantized image of Lincoln’s face may be analysed in fine to coarse levels of resolution. Even though different outputs resulting from different spatial-scale analyses may show interesting perceptual effects, the output resulting from a coarse analysis possesses its own global properties (such as global segmentation of the input based on physical connectivity) as well as its own local properties (such as orientation with each connected line segment), and the output from a fine analysis, its own local properties (such as orientation with each thinner line segment) as well as global properties (such as global segmentation also based on physical connectivity but with higher spatial resolution). Particularly, no claim can be simply made about precedence relations between the coarse and fine levels of resolution in their perceptual outputs. The contrast of spatial scales seems not to be essentially relevant to the “global and local”.

**“Global and local” vs. “whole and part”.** It is even more likely that to confuse “whole and part” with “global and local”, as typically shown by the famous “global precedence hypothesis” by Navon (1977), with its link to the “forest before trees”. The relation of “whole vs. part” simply implies the relation of “large-scale vs. small-scale”, as a whole is certainly larger than its parts. However, it is not difficult to see that no claim can be safely made that a whole is prior to its parts in perception. There is no reason to state that a forest, as a whole, would be certainly perceived prior to a tree, as a part of the forest. It obviously depends on viewing conditions. One more example is as follows. There is no reason to claim that a face, as a whole, would be perceived earlier than an eye or a mouth, as a part of the face. It obviously also depends on the conditions of looking. However, we are able to make a claim that the holistic registration, which determines which two eyes and which mouth belong to the same face, takes place before the percept of the face. That is, the perception of global topological invariant of connectivity of the face (or a connected component as a face) is prior to the perception of local properties of the face, such as the size and the shape of the face. Otherwise there would be no meaningful face recognition. The phenomena of “the global precedence” found with compound letters or figures are due to the fact that visual processes are “from global to local”, that is, in the case of grouping, grouping based on proximity is prior to grouping based on similarity, and grouping based on similarity of topological properties is prior to grouping based on similarity of other local geometric properties. Given the present definition of “global” and “local”, we can claim

“global precedence” in a precise way, that is, in the functional hierarchy of form perception, global topological perception is prior to perception of other local geometric properties. There is no such causality or time dependence as “whole precedence” or “large precedence”. The hierarchical structure of whole objects being composed of parts may be considered a natural aspect of the structure of the objective world. But this seems not to be a well-formed way to pose the psychological question about the priority of perceptual processes or about where the visual processing begins. This, I believe, is actually the deep cause why the experimental results on global precedence are very controversial.

*Global perception is prior to local perception: To reformulate holism*

As is well known, “The whole is different from the sum of its parts” is commonly quoted as the main point of holism, the central believe of Gestalt psychology. Certainly the distinction Gestaltists make between the whole and the sum of its parts is not new, and various fields, such as physics, chemistry, and information sciences, also make the similar claim. What is then the core and substantial contribution from Gestalt psychology to the understanding of the part-whole relationship? I believe the rich empirical observations about the whole-part relationship from Gestalt psychology suggest a logical priority and a priority of time dependence of “wholes”, that is, wholes are perceived directly without prior perception of and before the perception of its parts. However, the lack of precise and scientific understanding of the intuitive concept of “wholes” has hindered holistic approach.

From the perspective of the topological approach, we may state that “Global perception is *prior* to local perception” or “Holistic (or global) registration is *prior* to local analysis”, which goes beyond the notion that “The whole is different from than the sum of its parts”. Thus, in addressing the fundamental question of “Where visual processing begins?”, the most important scientific value contributed by Gestalt psychology lies in the line of thinking that global perception is *prior* to local perception, or holistic registration is *prior* to local analysis. It turns out that “prior” is the key word. In the precise and strict sense, “prior” implies two aspects of meaning. First, with respect to the time dependence, global perception occurs earlier than local perception, (particularly global topological perception based on physical connectivity occurs earlier than the perception of other geometrical properties). Second, with respect to causality, local perception depends on global organization—particularly, the perception of local geometrical properties depends on global topological perception. The time dependence of global topological perception and other local geometric properties was demonstrated by the functional hierarchy in form perception discussed in Part IV. The causality between topological perception

and local geometrical perception is clearly enlightened by Lappin's (1985, pp. 75–76) following arguments:

A fundamental problem in constructing a global geometric representation of a projected optical pattern is topological—partitioning the data into subsets of connected points associated with the same structure characterized by the same or smoothly changing parameter values. The measurement problem cannot be isolated from the connectedness problem. Measures of local geometric relations depend on the global organization of connections in space and time.

This line of thinking, that “global perception is prior to the local perception”, provides a formal expression of classic Gestaltist intuitions that the perception is holistic at its earliest levels, and therefore serves as a key to solve various major issues in perceptual organization. As we will see in the following discussion, this reformulation of Gestalt holism also plays a central role in clarifying the notion of direct perception, and rethinking the foundations of computational approaches to vision.

### Gibsonian psychology

As is well known, Gibson's core idea may be summarized as “direct perception of invariance”. As Gibson (1979, p. 249) emphasized, “The perceptual system simply extracts the invariants from the flowing array; it resonates to the invariant structure or is attuned to it.” In fact, the idea that stressed the important role of transformations and invariants preserved under transformation in perception is rooted in holism, as emphasized by Gestalt psychology (Palmer, 1999). While it shares the deep and reasonable insight of holism with Gestalt psychology, Gibsonian psychology, like Gestalt psychology, also suffered from the lack of a formal and precise theoretical treatment of its basic concepts. As Gibson had never been very clear about the nature of direct perception of invariance and the processes which mediate it, we should be more explicit to explain exactly what “invariance” means, exactly what “direct” means, and then precisely what is the implication of “direct perception of invariance”.

#### *The key step towards a more general understanding of invariance: Invariants preserved over shape-changing transformations*

Current approaches driven by the insight of invariance perception have made important contributions to our understanding of visual perception. Questions related to invariance perception, such as what kinds of invariant properties have psychological reality, are central to the contemporary study of visual perception. As emphasized earlier, invariants preserved over the shape-changing transformations are considered to play a fundamental role in introducing the topological



approach to perceptual organization. To address the relation of the topological approach to Gibson's theory of direct perception of invariance, it is natural to consider the relation of invariants under the shape-changing transformations to the invariants in Gibson's theory.

For illustrating the profound implication of shape-changing transformations, we will discuss briefly some of the current theories of invariance perception based on some specific transformation groups, for example, the symmetrical transformation group (e.g., Palmer, 1983, 1989, 1991), the affine transformation group (e.g., Norman & Todd, 1992; Tittle, Todd, Perotti, & Norman, 1995; Todd, 1998; Todd & Bressan, 1990; Todd & Norman, 1991; Zhang, 1994), and the projective transformation group (e.g., Cutting, 1986, 1998; Koenderink, van Doorn, Kappers, & Todd, 2000; Zhang & Chen, 2005). Studies with these invariance approaches have investigated which invariants, among various candidates, are perceptually useful in visual perception, with projective, affine, and symmetrical invariants being investigated in great detail.

For example, Todd and his co-workers (e.g., Norman & Todd, 1993; Tittle et al., 1995; Todd, 1998; Todd & Bressan, 1990; Todd & Norman, 1991) studied the visual perception of 3-D form from motion, from the perspective of the transformations of optical stimulation that occur when objects are observed in motion. Particularly, they found that observers were able, from the available information within two-frame motion sequences, to detect most types of non-rigid deformations, and to accurately discriminate structural properties that are invariant over affine transformations. But the observers had considerably more difficulty on tasks that require an accurate perception of Euclidean metric properties. Their finding, that observers can misperceive an object's extension in depth while correctly identifying that it is undergoing rigid rotation, suggests that Euclidean metric distances in three-dimensional space are not a primary perceptual component. It follows that human visual perception of structure from motion is essentially different from the early feature analysis, the primitives of which are typically local features based on Euclidean metric distances.

These theories, rooted in the idea of invariant perception, have provided some coherent frameworks for unifying interpretations of a range of phenomena in perceptual organization, and thus demonstrate the powerful role of invariants over a specific transformation in understanding of visual perception. However, a further question is whether Euclidean similarity, affine, and projective transformations are broad enough to cover the ranges of transformations that Gibson considered. As a matter of fact, for example, Cutting (1998) had already realized that the problem that the projective invariant of the cross-ratios was occasionally but not universally used by the visual system may be due to the invalidity of the rigidity assumption, namely, that perceivers often see things as nonrigid when they are rigid even when a rigid interpretation is possible.

To address this question, let us return to examine what kinds of transformations Gibson was originally concerned with. Gibson (1979) listed the eco-

logically meaningful changes of layout as follows: Rigid transformations and rotations of an object, nonrigid deformations of an object, surface deformations, and surface disruptions. In this tabulation, it is obvious that shape-changing transformations, particularly deformations (including the nonrigid deformations, surface deformations, and surface disruptions) go far beyond the projective, the affine, and the symmetrical transformations but were Gibson's major concern for the invariants in ecological optics. The current theories have mainly focused on more specific transformations but failed to pay attention to more general shape-changing transformations. This neglect of the more general transformations is costly. As emphasized before, the analysis of shape-changing transformations serves as a key for understanding various fundamental issues in perceptual organization. The failure to take an account of shape-changing transformations results in the loss of a more general and unified framework for understanding of invariance perception. This is particularly unfortunate, because Gibson himself had already knocked at the door of the more general transformations and more primitive invariants, but stopped just in front of it for the following reasons that we will discuss in details.

*A key for solving the problem of the one-to-one mapping assumption with precisely defining invariants: Global tolerance (topological) properties*

Even though he listed various changes in ecological environment, Gibson (1979) was nonetheless unsatisfied by a mere list of the various phenomena of invariants, in his own words (p. 310): "What is lacking is a theory of the invariants that preserved under disturbances with *nonpersistence* of units." "Can these disturbances of structure be treated mathematically?", Gibson asked, and answered: "It would simplify matters if all of these kinds of changes in the optic array could be understood as transformations in the sense of mappings, borrowing the term from projective geometry and topology" (p. 310). However, Gibson further commented, "But, unhappily, some of these changes *cannot* be understood as one-to-one mappings, either projective or topological" (p. 310). Gibson was correct in realizing the deep problem with applying the mathematical terms of projective geometry and topology to describe invariants in ecological environment: While mathematically topological and projective mappings require a fundamental assumption of one-to-one transformations, ecological changes don't always conform to this assumption, inasmuch as the array gains or loses units in time (p. 108). By acknowledging this mathematical difficulty in applying the topological approach to invariance perception Gibson demonstrated a deep insight into the nature of ecological change. However, unfortunately he failed to find a mathematical way out.

The definition of the topological transformation group in general topology involves two underlying assumptions of "one-to-one" and "continuous"

mappings. In contrast to the essentially local nature of “continuous” and “one-to-one” mappings, the mathematical structure of tolerance spaces is well adapted to ignore local variations for capturing global properties. In Part III, we faced the fundamental question of “How to define global properties in a discrete set”, which appeared to contradict the underlying assumption of the “continuous” mapping. The problem with the discrete nature of perceptual organization was solved by the mathematics of tolerance spaces in Part III. To apply topological invariants to describe invariants in ecological optics, however, appears to contradict directly another underlying assumption of the “one-to-one” mapping. One of the most powerful and important mathematical properties of tolerance spaces is that, from its definition, a tolerance mapping, more precisely, a tolerance homeomorphism doesn’t require the assumption of “one-to-one” mapping. That is, a tolerance homeomorphism does not need to meet the assumption of “one-to-one” mappings but only the assumption of so called tolerance embedding: A mapping of  $f$  is a tolerance embedding if the following condition holds:  $x_1 \sim x_2$  in  $\xi$  if and only if  $f(x_1) \sim f(x_2)$  in  $\eta$ . Let us take an example from Zeeman (1965) to illustrate the idea of how and why the assumption of “one-to-one” mapping is not necessary in a topological (homology) theory of tolerance spaces. Suppose  $(X, \xi)$  is a tolerance space. Let  $X$  be the finite set of atoms in the paper of which this page is made, and let  $\xi$  be the tolerance given by  $x_1 \sim x_2$  if  $x_1$  and  $x_2$  are less than, for example, a millimetre apart. Suppose  $(Y, \eta)$  to be another tolerance space. Let  $Y$  be the (infinite) set of points on this page, and  $\eta$  be the tolerance given by  $y_1 \sim y_2$  if  $y_1$  and  $y_2$  are less than a millimetre apart. Let  $f: X \rightarrow Y$  be the function that assigns to each atom the nearest point on the surface of the paper. Then  $f$  is a tolerance homeomorphism. This example illustrates the very powerful idea that a finite set  $X$  can be homeomorphic to an infinite set  $Y$ , to within some tolerance. As Zeeman (1965) argued: “the idea lies at the root of the whole theory, because our brains, which are finite in quality, have the power to handle continuous information, which is infinite in quality, to within some tolerance”. Importantly, under the tolerance homeomorphism, global tolerance (topological) properties, such as the number of tolerance holes, tolerance connectivity, and dimension, are preserved (Zeeman, 1965).

As shown in Part III regarding the fundamental question of how to define global properties in a discrete set, we see once again the important power of tolerance spaces to handle various issues related to the global nature of perceptual organization. With the mathematics of tolerance spaces there is a way out for Gibson’s efforts. It is the global characteristic of tolerance spaces that provides a suitable mathematical treatment on ecological invariance perception. After removing the essentially local assumption of “one-to-one” mapping, we are able to claim that the invariant structures in optical flow in Gibson’s theory may be properly described as invariant properties at different levels of geometry, particularly at the level of topology.

*Direct perception: The priority of holistic (global) perception*

Let us turn to another and the most controversial aspect of Gibson's theory, that is, "direct perception". The concept of "direct" is also essentially a holistic concept, contrasting with the early feature analysis. It is, therefore, interesting to discuss the relation between the global topological perception and Gibson's idea of direct perception.

*Against "against direct perception".* A major argument "against direct perception" (Ullman, 1980) was originally from Marr (1982, p. 30): "Although one can criticise certain shortcomings in the quality of Gibson's analysis, its major, and, in my view, fatal shortcoming lies at a deeper level and results from a failure to realize two things. First, the detection of physical invariants, like image surfaces, is exactly and precisely an information-processing problem, in modern terminology. And, second, he vastly underrated the sheer difficulty of such detection." As pointed out in Part I, in the eyes of Marr, the question—what are the primitives of visual representation at early stage of vision—is not an empirical issue to be answered by experimental studies, but essentially a theoretical issue to be answered by computational complexity analyses regarding "what it is possible to compute" (Marr, 1978). It is true that it is difficult to create computational models that will actually extract invariants, particularly topological invariants. However, the difficulty of computing invariants should not be considered as sufficient reason against direct perception of invariants. As emphasized at the beginning of the present paper, physically or computationally simple does not necessarily mean psychologically simple or perceptually primitive; therefore, the question of which variables are perceptual primitives is not a question answered primarily by logical reasoning or computational complexity analysis but rather empirical finding. In particular, even though the global nature of topological properties makes their computation difficult, the data reviewed here demonstrate that topological perception is prior to the perception of local geometrical properties. In the functional hierarchy of form perception, the order of priorities of perceiving topological properties and other local geometrical properties is just reversed in comparison with the order of difficulties in their computations. The "higher order invariants" may be computationally higher order but psychologically primitive.

*Direct perception and priority of global perception over local feature analysis.* Direct perception can be understood as the priority of global perception over local feature analysis. Here priority has the exact two senses as emphasized before, with respect to time dependence, and with respect to causality. First, with respect to time dependence, the perception of global invariants (particularly topological invariants) is direct without the mediating processes of local feature analysis. From the time dependence of global priority,

holes are perceived directly, not by adding up a set of features like edges or blobs, which would be decomposed before the global perception. The perception of global topological invariants of an object (such as connectivity and holes) is direct, without mediating feature analysis of the object, because it occurs before perceiving local features of the object.

Second, with respect to the causal relationship, global invariants, particularly topological invariants, are perceived directly without the prior need of local feature analysis. The causality is just the opposite of the early feature analysis: The measure of local geometrical properties depends on global topological perception, that is, it depends on the global organization of connections in space and time. For perceiving global properties of an object it is not necessary, or even possible, to analyse the object into its featural properties. The local featural analysis must be based on the global topological organization, namely, segmentation of the raw image input into potential objects, on which measures of local geometric properties and relations are built up. In this way, the perceptual system simply extracts the topological invariants; it resonates directly to the invariant structure.

To make the above discussion more illustrative, let us appeal to an example concerning the perception of a hole for a triangle stimulus. The visual system achieves direct perception of the hole of the triangle, rather than that first analysing the triangle into three line segments composing the triangle, and then binding them together to form the integrated percept of the triangle with a hole. The indirect process looks reasonable as a consequence of the early feature-analysis theory, but is contradicted by the convergent experimental data summarized in this paper.

In summary, concerning the nature of the input for vision, Gibson's theory and the topological approach share the viewpoint that the starting input for visual perception should be considered as "higher order"<sup>2</sup> invariants rather than local features with simple components. Moreover, "direct perception" may be understood from the perspective of global priority: Global topological perception is direct in the sense that it occurs earlier than and determines the local feature analysis.

### Early topological perception: A challenge to computational approaches to vision

It is well known that Minsky and Papert's classic results on the high complexity of topological properties, in their milestone book *Perceptrons* (1969), have

---

<sup>2</sup> The term of "higher order" was also used by Gibson and his followers. However, even though they are computationally or physically "high order", topological invariants are perceptually or ecologically primitive.

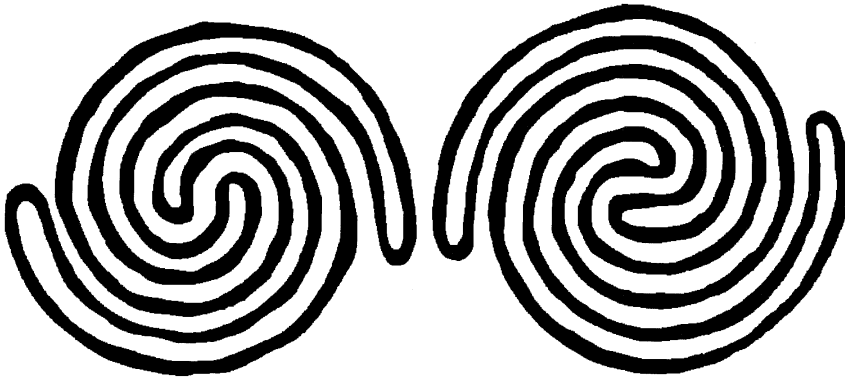
created much controversy about the capabilities and limitations of connectionist models. *Perceptrons* was blamed for causing a “dark age” for connectionism throughout the 1970s. Actually, the most influential mathematical result of the book is of the limitation of perceptrons for computing topological properties. That is, for perceptrons, the topological predicate is not finite order; however lower order perceptrons can compute local geometrical properties. The question raised by topological computation has thus been one of issues that lie at the core debate between traditional AI and connectionism. Human topological perception, and its relation to topological computation, therefore, has inevitably raised some interesting issues about the foundations of computational approaches to vision.

*Scepticism towards human topological perception is rooted in computational presumption on vision*

Proponents of computational approaches to vision are sceptical of the topological hypothesis because, in their words, “The determination of topological properties is not likely to be primitive for any kind of computational system” (Rubin & Kanwisher, 1985). The very nature of the computational approach to vision makes them assume that primitive elements of computation are simple components with local geometrical features, typically line segments and blobs. As is well known, in Marr’s computational system, the visual representation is built on the “primal sketch”, which emphasizes “the local geometry of an image”; the visual processing begins with computing the primal sketch, such as oriented edges, lines, and blobs, and all subsequent computations are implemented as manipulations of the primal sketch. How should we consider the relationship between the early topological perception and the primal sketch?

More recently, connectionists have held two distinctive positions about the relationship between topological perception and topological computation. On the one hand, as Anderson and Rosenfeld (1988, p. 159) assumed, perceptrons act as psychological models, and considered this point to be what “became quite important in the resurrection of neural networks”. Their arguments were largely based on the famous demonstration of the spirals (Figure 31, from Minsky & Papert, 1969, Fig. 5.1). The phenomenon that, even though one of the two spirals is physically connected and the other not, we are unable to tell this, was used to argue that like simple perceptrons cannot compute the predicate connectedness, “we cannot compute connectedness either” and “we act as if we had some of the same limitations that a perceptron does” (Anderson & Rosenfeld, 1988, p. 159).

It is even more interesting to note that Rumelhart and McClelland (1986) apply back propagation and Boltzmann machines to argue that Minsky and Papert’s (1969) result about topological computation is not valid in general. They considered that Minsky and Papert’s theorem is applicable only to simple



**Figure 31.** The spirals as an ambiguous figure in connectedness (from Minsky & Papert, 1969, Figure 5.1). Phenomenally, even though the left one of the two spirals is physically connected and the right one not, we are unable to tell this.

one-layer perceptrons but not to the larger class of perceptron-like models, such as multilayered perceptrons and Boltzmann machines, and it is not difficult to develop networks capable of computing topological properties, such as connectedness or the inside/outside relationship (Rumelhart & McClelland, 1986). There is no doubt that connectionist models, such as back propagation and Boltzmann machine, are of great interest in providing insights into topological computation. However, can multilayered networks or Boltzmann machines really serve as a computational model for human topological perception? The following addresses not only (1) the scepticism towards humans' ability to perceive topological properties, but also (2) the validity of Minsky and Papert's classic results on the high complexity of topological computations.

#### *Inadequacy of scepticism towards topological perception*

A primary problem in the study of topological perception is that mathematical reasoning is often confused with psychological consideration. It is worth emphasizing again that apparent counterexamples against topological perception often result from the neglect of differentiating two pairs of concepts: Visibility vs. difficulty of discrimination, and physical (spatial) connectedness vs. other organizational factors (such as similarity). As these factors were differentiated rather than neglected, the counterexamples were completely ruled out (e.g., Chen, 1990).

The two considerations are also applicable to the counterexample of the famous spirals (Figure 31), where the connectivity appears to be undetectable by our visual system; not only physical connectedness but also other organizational

factors such as proximity (of neighbouring curves) and similarity (of grey-level, width, curvature) are present in the spirals. There are competing organizations produced by the several simultaneous factors. In addition, when the central part of the spiral is fixed, the reduced visibility of its surrounding part may not provide a clear clue about physical connectedness. In this case, the competing organizations may dominate grouping so that alternative percept of the spiral's connectedness occurs. When these additional factors, other than physical connectedness, are considered, then, I suggest, there is no support for the scepticism that "we cannot compute connectedness either" (Anderson & Rosenfeld, 1988).

The scepticism towards topological perception actually implies that the topological perception may be unnecessary for visual function. However, the above discussion about the causality between topological perception and local geometrical perception indicates that topological perception is necessary for visual function, because perception of any local geometric properties depends on topological perception, namely, the global organization of connections in space and time.

*Early topological perception: A challenge to computational approaches to vision based on local features*



Connectionist approaches to vision, like classic computational approaches, tend to assume local geometrical features to be primitives. Existing connectionist models, therefore, lead one to expect that discriminations based on local geometrical properties would occur earlier than those based on topological properties. For example, consider a connectionist model by Sejnowski and Hinton (1987) for segregating figure and background by computing inside/outside relation. It is interesting to note that as they proved, in comparison with computing stereopsis, computing inside/outside relation is particularly difficult: While the search space of the stereo algorithm is fairly well behaved, the global minimum is shallower and the search space has many local minima in the search space of computing the inside/outside relation. This model also shares the same approach as other computational approaches to vision, in assuming local geometrical features to be primitive elements of computation. In their model, "The 'bottom-up' input is not the raw image itself but is highly processed version of the image containing the location and orientation of edges" (Sejnowski & Hinton, 1987). Hence the computation of inside/outside relation is implemented as manipulations of these primitive elements. This leads one to expect that discrimination based on local geometrical properties would occur earlier than that based on topological properties—exactly the opposite to the empirical results summarized above to show *early* topological perception.

The important development of back propagation has also led to doubts as to whether Minsky and Papert's (1969) classic result on the high complexity of



topological properties is broad enough to be true for multilayered networks (Rumelhart & McClelland, 1986). In my opinion, since Minsky and Papert's result is about complexity rather than computability, it is not sufficient to question the validity and generality of their result only by developing a sample of network capable of computing topological properties under certain limited conditions. It might even be possible to construct some samples of multilayered networks someday that are capable of computing in real time some aspects of topological structure under certain limited conditions. However, in comparison with other geometrical properties, topological properties still have higher complexity even for multilayered networks. As a consequence, the difficulty remains for computational approaches based on local geometrical features to explain *early* topological perception. As Cutting (1987) pointed out, topology is "the only kind of geometric information not ultimately couched as visual angles". So, it is not a surprise at all that Minsky and Papert elegantly proved that all topologically invariant predicates (except one for Euler number) are not finite order. After Minsky and Papert, some different measures of complexity of geometrical properties and some extensions of their complexity measure have been advanced. However, connectedness still turns out to be highly complex (e.g., Rosenfeld & Kak, 1982). Minsky and Papert (1988) argued that multilayered networks with back propagation could compute some topological properties, but when the sample of input increases, the number of hidden units needed and the time to learn will increase tremendously. More recently, Wang (2000) proved: "For two-dimensional (2D) R, the number of connected patterns grows exponentially with respect to  $|R|$ . Thus, for R not too small, a practical training process can employ only a tiny percentage of all possible training samples, and it is hard to expect that a multilayer perceptron can successfully generalize from relatively so few training samples", and he further argued: "Given this consideration and that no success has been reported on recognizing connectedness, it is reasonable to project, based on the result on simple perceptrons, that the limitation exists for multiplayer perceptrons." Topological computation is still a major difficulty to connectionism (Chen, 1989).

This reminds us again of the ancient Chinese proverb cited at the outset of this paper: Everything is difficult at its very beginning. The topological approach may be merely a first step in the long march to understand vision, but importantly, I believe, a step in the correct direction.

## REFERENCES

- Anderson, J. A., & Rosenfeld, E. (1988). *Neurocomputing*. Cambridge, MA: MIT Press.
- Anstis, S. (1986). Motion perception in the frontal plane. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory process and perception* (pp. 16–11–27). New York: Wiley.

- Beck, J. (1967). Perceptual grouping produced by line figures. *Perception and Psychophysics*, 2, 491–495.
- Beck, J. (1972). Similarity grouping and peripheral discriminability under uncertainty. *American Journal of Psychology*, 85, 1–19.
- Beck, J., & Ambler, B. (1973). The effects of concentrated and distributed attention on peripheral acuity. *Perception and Psychophysics*, 14, 225–230.
- Chen, L. (1982). Topological structure in visual perception. *Science*, 218, 699–700.
- Chen, L. (1983). What are the units of figure perceptual representation? *Studies in Cognitive Sciences* (No. 22). School of Social Science, University of California, Irvine, CA, USA.
- Chen, L. (1985). Topological structure in the perception of apparent motion. *Perception*, 14, 197–208.
- Chen, L. (1986). Topological perception: A possible dark cloud over computational approaches to vision. In X. S. Qian (Ed.), *On cognitive science*. Shanghai: Shanghai People's Press.
- Chen, L. (1989). Topological perception: A challenge to computational approaches to vision. In R. Pfeifer (Ed.), *Connectionism in perspective*. North-Holland: Elsevier Science Publishers.
- Chen, L. (1990). Holes and wholes: A reply to Rubin and Kanwisher. *Perception and Psychophysics*, 47, 47–53.
- Chen, L. (2001). Perceptual organization: To reverse back the inverted (upside down) down question of feature binding. *Visual Cognition*, 8, 287–303.
- Chen, L., Zhang, S. W., & Srinivasan, M. (2003). Global perception in small brains: Topological pattern recognition in honeybees. *Proceedings of the National Academy of Sciences*, 100, 6884–6889.
- Chen, L., & Zhou, W. (1997). Holes in illusory conjunctions. *Psychonomic Bulletin and Review*, 4, 507–511.
- Cutting, J. E. (1986). *Perception with an eye for motion*. Cambridge, MA: MIT Press.
- Cutting, J. E. (1987). Perception and information. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 38, pp. 61–90). Palo Alto, CA: Annual Review Inc.
- Cutting, J. E. (1998). Information from the world around us. In J. Hochberg (Ed.), *Perception and cognition at century's end*. New York: Academic Press.
- Desimone, R., & Duncan, J. (1995). Neural mechanism of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501–517.
- Duncan, J. (1989). Parallel processing: Giving up without fight. *Behavioral and Brain Sciences*, 12, 402–403.
- Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Flavell, J. H., & Draguns, J. (1957). A microgenetic approach to perception and thought. *Psychological Bulletin*, 54, 197–217.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Han, S., Humphreys, G., & Chen, L. (1999a). Parallel and competitive processes in hierarchical analysis: Perceptual grouping and encoding of closure. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1411–1432.
- Han, S., Humphreys, G., & Chen, L. (1999b). Uniform connectedness and classical Gestalt principles of perceptual grouping. *Perception and Psychophysics*, 61, 661–674.
- Hubel, D. H., & Weisel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28, 229–289.
- Hubel, D. H., & Weisel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195, 215–243.
- Hughes, H. C., Norzawa, G., & Kitterle, F. (1996). Global precedence, spatial frequency channels, and the statistics of natural images. *Journal of Cognitive Neuroscience*, 8, 187–230.

- Humphreys, G. W. (2001). A multi-stage account of bidding in vision: Neuropsychological evidence. *Visual Cognition*, 8, 381–410.
- Humphreys, G. W., & Bruce V. (1989). *Visual cognition: Computational, experimental and neuropsychological perspectives*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Humphreys, G. W., Romani, C., Olson, A., Riddoch, M. J., & Duncan, J. (1994). Non-spatial extinction following lesion of the parietal lobe in Humans. *Nature*, 372, 357–359.
- Julesz, B. (1965, February). Texture and visual perception. *Scientific American*, 212, 38–48.
- Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, 290, 91–97.
- Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Kinchla, R. A., & Wolfe, J. M. (1979). The order of visual processing: “Top-down,” “bottom-up,” or “middle-out”? *Perception and Psychophysics*, 25, 225–231.
- Koenderink, J. J., van Doorn, A. J., Kappers, A. M. L., & Todd, T. (2002). Pappus in optical space. *Perception & Psychophysics*, 64, 380–391.
- Kolers, P. A., & Pomerantz, J. R. (1971). Figural change in apparent motion. *Journal of Experimental Psychology*, 87, 99–108.
- Kubovy, M., Holcombe, A. O., & Wagemans, J. (1998). On the lawfulness of grouping by proximity. *Cognitive Psychology*, 35, 71–98.
- Kubovy, M., & Pomerantz, J. (1981). *Perceptual organization* (pp. 425–456). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- LaGasse, L. (1993). Effects of good form and spatial frequency on global precedence. *Perception and Psychophysics*, 53, 89–105.
- Lamb, M. R., & Yund, E. W. (1996). Spatial frequency and attention: Effect of level-, target-, and location-repetition on the processing of global and local forms. *Perception and Psychophysics*, 58, 362–373.
- Lappin, J. S. (1985). Reflection on Gunnar Johansson’s perspective on the visual measurement of space and time. In W. H. Warren, Jr., & R. E. Shaw (Eds.), *Persistence and changes: Proceedings of the first international conference on event perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Marr, D. (1978). Representing visual information: A computational approach. *Lectures on Mathematics in the Life Sciences*, 10, 61–80.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman.
- Minsky, M. L., & Papert, S. (1969). *Perceptrons: An introduction to computational geometry*. Cambridge, MA: MIT Press.
- Minsky, M. L., & Papert, S. (1988). *Perceptrons: An introduction to computational geometry* (Exp. ed.). Cambridge, MA: MIT Press.
- Müller, H. J., Elliott, M. A., Herman, C. S., & Mecklinger, A. (2001). Introduction. *Visual Cognition*, 8, 273–285.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353–383.
- Norman, J. F., & Todd, J. T. (1993) The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception and Psychophysics*, 3, 279–291.
- Olson, R. K., & Attneave, F. (1970). What variables produce similarity grouping? *American Journal of Psychology*, 83, 1–21.
- Palmer, S. (1983). The psychology of perceptual organization: A transformation approach. In J. Beck, B. Hope, & A. Baddeley (Eds.), *Human and machine vision* (pp. 269–339). New York: Academic Press.

- Palmer, S. (1989). Reference frames in the perception of shape and orientation. In B. E. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 121–163). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Palmer, S. (1991). Goodness, Gestalt, groups, and Garner: Local symmetry subgroups as a theory of figural goodness. In G. R. Lockhead & S. J. R. Pomerantz (Eds.), *The perception of structure: Essays in honor of Wendell R. Garner* (pp. 23–39). Washington DC: American Psychological Association.
- Palmer, P. (1999). *Vision science*. Cambridge, MA: MIT Press.
- Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin and Review*, 1, 29–55.
- Pomerantz, J. R. (1978). Are complex visual features derived from simple ones? In E. L. J. Leeuwenberg & H. F. J. M. Buffart (Eds.), *Formal theories of visual perception* (pp. 217–229). Chichester, UK: Wiley.
- Pomerantz, J. R. (1981). Perceptual organization in information processing. In M. Kubovy & J. Pomerantz (Eds.), *Perceptual organization* (pp. 141–180). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Pomerantz, J. R., Sager, L. C., & Stoever, R. J. (1977). Perception of wholes and of their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 422–435.
- Pöppel, E. (1985). *Mind works: Time and conscious experience* (G. des Bewusstseins, Trans.). Orlando, FL: Harcourt Brace Jovanovich Publishers.
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Science*, 1, 56–61.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for parallel tracking mechanism. *Spatial Vision*, 3, 179–197.
- Ramachandran, V. S., Ginsburg, A. P., & Anstis, S. M. (1983). Low spatial frequencies dominate apparent motion. *Perception*, 12, 457–461.
- Rees, G., Friston, K., & Koch, C. (2000). A direct quantitative relationship between the functional properties of human and macaque V5. *Nature Neuroscience*, 3(7), 716–723.
- Rock, I. (1986). The description and analysis of object and event perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2. Cognitive processes and performance* (pp. 1–46). New York: Wiley.
- Rock, I., & Palmer, S. (1990, December). The legacy of Gestalt psychology. *Scientific American*, 48–61.
- Rosenfeld, A., & Kak, A. (1982). *Digital picture processing 2*. New York: Academic Press.
- Rubin, J. M., & Kanwisher, N. (1985). Topological perception: Holes in an experiment. *Perception and Psychophysics*, 37, 179–180.
- Rumelhart, D. E., McClelland, J. L., & The PDP Group. (1986). *Parallel distributed processing* (Vol. 1). Cambridge, MA: MIT Press.
- Sejnowski, T. J., & Hinton, G. E. (1987). Separating figure from ground with a Boltzmann machine. In M. A. Arbib & A. R. Hanson (Eds.), *Vision, brain and cognition* (pp. 703–724). Cambridge, MA: MIT Press.
- Srinivasan, M., & Zhang, S. W. (1998). Probing perception in miniature brain: Pattern recognition and maze navigation in honeybees. *Zoology*, 101, 246–259.
- Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1995). The systematic distortion of perceived 3D structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 663–678.
- Todd, J. T. (1998). Theoretical and biological limitations on the visual perception of three-dimensional structure from motion. In T. Watanabe (Ed.), *High-level motion processing: Computational, neurophysiological and psychophysical perspectives* (pp. 359–380). Cambridge MA: MIT Press.

- Todd, J. T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception and Psychophysics*, 48, 419–430.
- Todd, J., Chen, L., & Norman, F. (1998). On the relative salience of Euclidean, affine, and topological structure for 3-D form discrimination. *Perception*, 27, 273–282.
- Todd, J. T., & Norman, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception and Psychophysics*, 50, 509–523.
- Treisman, A. (1985). Features and objects in visual processing. *Scientific American*, 255, 114–125.
- Treisman, A. (1986). Properties, parts, and objects. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2. Cognitive processes and performance* (pp. 1–70). New York: Wiley.
- Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, 40A, 201–237.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15–48.
- Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- Ullman, S. (1980). Against direct perception. *Behavioral & Brain Sciences*, 3, 373–415.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. G. Ingle, M. A. Goodale, & R. J. Q. Mansfield (Eds.), *Analysis of visual behaviour* (pp. 549–586). Cambridge, MA: MIT Press.
- VanMarle, K., & Scholl, B. J. (2003). Attentive tracking of objects vs. substance. *Psychological Science*, 14, 498–504.
- Wang, D. L. (2000). On connectedness: A solution based on oscillatory correlation. *Neural Computation*, 12, 131–139.
- Wehner, R. (1971). The generalization of directional visual stimuli in the honey bees, *Apis mellifera*. *Journal of Insect Physiology*, 17, 1579–1591.
- Weisstein, N., & Harris, C. S. (1974). Visual detection of line-segments: An object-superiority effect. *Science*, 186, 752–755.
- Wolfe, J. M. (1994). Guided Search 2.0: A revised model of guided search. *Psychonomic Bulletin and Review*, 1, 202–238.
- Zeeman, E. C. (1962). The topology of brain and perception. In M. K. Fort (Ed.), *Topology of 3-manifolds and related topics* (pp. 240–256). Englewood Cliffs, NJ: Prentice Hall.
- Zeeman, E. C. (1965). Topology of the brain. In *Mathematics and computer science in biology and medicine: Proceedings of the conference held by the Medical Research Council in Association with the Health Department, Oxford, July 1964* (pp. 277–292). London: Her Majesty's Stationary Office.
- Zhang, J. (1994). Image representation using affine covariant coordinates. In Y.-L. O, A. Toet, D. Foster, H. J. A. M. Heijmans, & P. Meer (Eds.), *Shape in picture: Mathematical description of shape in grey-level images* (pp. 353–362). Berlin: Springer-Verlag.
- Zhang, J. (1995). Motion detectors and motion segregation. *Spatial Vision*, 9, 261–273.
- Zhang, J., & Chen, L. (2005). Characterizing projective geometry of binocular visual space through Moebius transformation. *Manuscript in preparation*.
- Zhang, J., & Wu, S. (1990). Structure of visual perception. *Proceedings of the National Academy of Sciences USA*, 87, 7819–7823.
- Zhuo, Y., Zhuo, T. G., Rao, H. Y., Wang, J. J., Meng, M., Chen, M., Zhou, C., & Chen, L. (2003). Contributions of the visual ventral pathway to long-range apparent motion. *Science*, 299, 417–420.

