Perceptual Use of Nonaccidental Properties

JOHAN WAGEMANS *Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium*

Abstract Under the assumption of a general viewpoint, particular image properties, such as cotermination, straightness, and parallelism, can be used to infer, more or less reliably, the corresponding characteristics in the world. In this paper, the literature about these nonaccidental properties (NAPs) is reviewed to trace its historical roots, to list the properties that function as NAPs, and to discuss the psychological evidence for their detection and use. Against this background, four experiments are reviewed and four are fully described that were designed to test the perceptual use of skewed symmetry (SS), which results from orthographic projection of planar bilateral or mirror symmetry (BS). Despite the large symmetry advantage obtained in all experiments, SS is only perceived as BS-in-depth in cases of closed polygons or dot patterns with higher-order types of symmetry. In all random dot patterns and in some symmetric patterns with low "Gestalt", subjects relied on more local groupings which are qualitatively affine invariant, such as clusters based on proximity or curvilinearity. Based on previous approaches in the literature and these new findings, I suggest some distinctions between different ways of using NAPs, which might foster further research.

Résumé Dans l'hypothèse où il existe un point de vue général, il est possible de déduire de façon plus ou moins juste, à partir de certaines propriétés illustrées dans les images comme la cotermination, la rectilignité et le parallélisme, les caractéristiques correspondantes dans le monde. Dans le présent document, on passe en revue la littérature concernant les propriétés non accidentelles (PNA) afin d'en exposer l'origine, d'énumérer les propriétés constituant des PNA et d'examiner les éléments de preuve psychologiques à l'appui de l'identification et de l'utilisation de ces propriétés. À la lumière de ces informations, quatre expériences sont examinées, et quatre autres, décrites de façon exhaustive; elles ont été conçues pour vérifier l'utilisation perceptuelle de la symétrie oblique, résultant de la projection orthographique de la symétrie bilatérale planaire ou spéculaire. Malgré l'important avantage de la symétrie qu'on a observé dans toutes les expériences, les sujets ne percevaient que de la symétrie bilatérale en profondeur là où il y avait symétrie oblique, dans le cas de polygones fermés ou de motifs de points dont l'ordre de

Canadian Journal of Psychology, 1992, 46:2, 236-279

symétrie était supérieur. Pour tous les motifs de points choisis au hasard et pour certains motifs symétriques possédant peu de «gestalt», les sujets se fiaient à des groupements plus locaux qui étaient affines et invariants sur le plan qualitatif, par exemple des amas fondés sur la proximité ou sur la curvilinéarité. Compte tenu des approches déjà utilisées et des résultats des expériences susmentionnées, je propose certaines distinctions entre les façons d'utiliser les PNA, qui pourraient conduire à d'autres recherches.

A basic problem in perceptual psychology and computer vision is the recovery of three-dimensional (3-D) information from two-dimensional (2-D) inputs, which is mathematically seriously underdetermined. For example, the two sets of curves in 3-D space (denoted by A and B), depicted in Figure 1 (taken from Lowe, 1985), project to identical curves in the 2-D image plane. Hence, it seems an impossible task for a human perceiver or a computer vision algorithm to derive the "real-world situation" from the available images. Yet, humans are very accurate in recognizing 3-D objects from line drawings of their contours only (e.g., Biederman & Ju, 1988). To explain how this is achieved in biological vision systems and to attempt to realize the same kind of success in artificial vision systems are formidable tasks for perceptual psychology and computer vision, respectively.

As proposed by Lowe (1985), a potentially fruitful way to start a specification of the problem and an attempt towards its solution is suggested by careful consideration of Figure 1. Although both depicted 3-D scenes project to the image in Figure 1, it is much more likely that the image results from the curves in A than from those in B. More specifically, it is quite unlikely that curve terminations which are separated in 3-space would project to a common point in the image, that a curved line would appear straight, or that nonparallel lines would appear parallel. In fact, there is only one particular viewpoint from which this could be the case. Under the assumption of a general viewpoint, particular image properties, such as cotermination, straightness, and parallelism, can be used to infer, more or less reliably, specific characteristics in the world. These image regularities are called *nonaccidental properties* (NAPs for short), because of the assumption that they do not accidentally result from a singular arrangement of objects and viewer (camera) in the scene.

Although inferences based on this kind of evidence do not yield perfect solutions in all cases (indeed, there is one viewpoint for which it does not work), the suggestion is that they are the kind of information the human visual system is relying on in solving the recovery problem. In addition, if it were possible to make an algorithm detect these properties sufficiently reliably and implement a kind of Bayesian inductive inference or evidential reasoning, this could be used to provide robots with similar capabilities of solving the

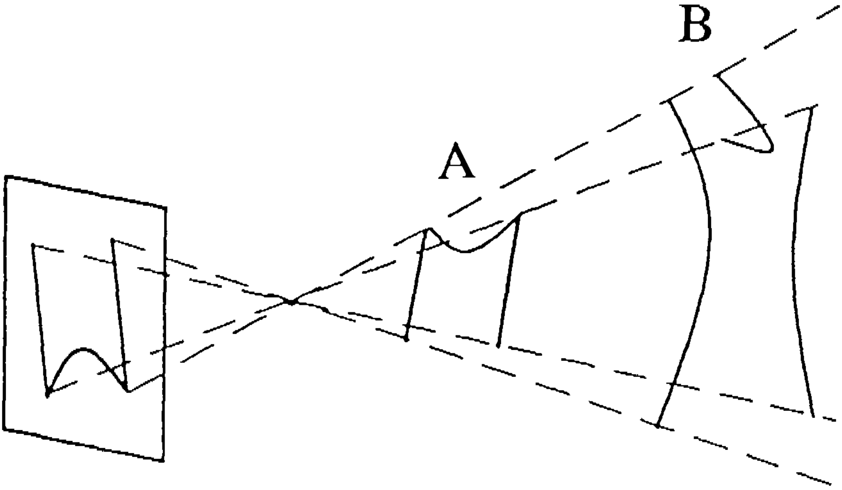

Fig. 1 Recovery problem and NAPs as potential solution. In principle, the curves in the image plane can result from an infinite number of sets of curves in the 3-D world, two of which are depicted. However, assuming that the coterminations and parallelism in the image are not an accidental result of a particular viewpoint, the curves in A are much more likely than those in B. From D. G. Lowe, 1985, *Perceptual Organization and Visual Recognition*, p. 74. Copyright 1985 by Kluwer.

recovery problem. Recent work by Lowe (1985, 1987a and b, 1989, 1990) shows that this approach is feasible indeed. Both subproblems, perceptual organization (to establish the required spatial relations such as parallelism) and evidential reasoning, constitute essential parts of his SCERPO system (i.e., an acronym for Spatial Correspondence, Evidential Reasoning, and Perceptual Organization). More detailed information about the way it works exactly (which is beyond the scope of this paper) can be found in Lowe's thesis (1985), some overviews (1987a, 1990), or in the papers specifically devoted to both critical issues (1989, 1987b, respectively).

This paper consists of four parts. In Part I, the literature on NAPs is reviewed to provide the necessary background for some experiments about the use of skewed symmetry (SS), which results from orthographic projection of planar bilateral or mirror symmetry (BS), as a NAP in human visual perception (presented in Parts II and III). In the final part, I will suggest some distinctions between different ways of using NAPs, based on the findings and theories discussed in all previous parts.

NAPs: Approaches and Evidence

HISTORY

Although the acknowledgement of NAPs as a cornerstone for object recognition is only a recent development (e.g., Biederman, 1987; Lowe, 1987a), the logic behind it was known for some time already. In fact, the classic demonstration with Ames' distorted room (Ittelson, 1952) is a good example that human perceivers can be "cheated" if situations are set up which result in accidental properties in the image when viewed from a particular position (i.e., peephole). A whole school of perceptual theories (e.g., probabilistic functionalism, transactional psychology) was based on this kind of demonstrations (e.g., Brunswik, 1952; Ittelson, 1960, resp.). More recently, some theorists have formulated more or less the same property of nonaccidentalness in different terms. For example, Shepard's (1981) principle of structural stability and Rock's (1983) principle of coincidence avoidance are also based on the assumption of nonaccidentalness.

With respect to computer vision work, directly relevant to the issue of nonaccidentalness, two previous contributions must be pointed out. First, in his doctoral thesis, Stevens (1980) proposed an insightful distinction between invariants and constraints and defended the use of the latter in computer vision. He stressed that only a few properties in the 3-D scene are invariant over the perspective projection onto the image, and that, of those that are, only a few have the necessary feature of having an invariant inverse, so that the presence of the relation or property in the image does not necessarily imply the corresponding scene property. For example, if two points are close together in space, they invariably appear so in the image, but 2-D proximity in the image does not guarantee close positions in 3-space.

More relevant to vision are those properties which are necessarily present in the 3-D scene, when present in the image, or whose presence in the image can at least be used to constrain the possibly corresponding 3-D states. In short, the property of being an invariant holds for a feature in the 3-D world when projected on a 2-D image, whereas the property of being a constraint holds for a feature in the 2-D image when recovering a 3-D world from it. In addition to formulating a list of invariant or noninvariant properties which can be used as constraints, Stevens (1980, p. 21) mentioned the nonaccidentalness assumption without giving it a name:

> "For example, while parallel edges in the image do not invariably correspond to parallel 3-D edges, in order for the parallelism to be misleading, there must be a particular arrangement between the viewer and the 3-D edges. If the a priori probability is low for this to occur, then image parallelism would be useful for inferring 3-D structure."

Kanade (1981, p. 422) was even more explicit in his formulation of what he called the meta-heuristic of nonaccidental regularities: "Regularities observable in the picture are not by accident, but are some projection of real regularities". In addition to specifying some examples falling under this meta-heuristic (such as parallelism and skewed symmetry), Kanade proposed techniques by which they could be used in recovering the 3-D shape of an object from a single view (i.e., *mapping of the image regularities into the constraints in gradient space*; see also Kanade & Kender, 1983).

PROPERTIES THAT FUNCTION AS NAPS

More important than tracing the roots of this approach is to see what kind of *image regularities can function as* NAPs *in the sense that they can be used to* reliably recover (or constrain the set of possible solutions in recovering) 3-D scene structure. Although not all theories agree in this respect, some characteristics that definitely do not fulfil the necessary requirements can be indicated (e.g., Biederman & Shiffrar, 1987). For example, metric properties such as length, angle, degree of curvature are not preserved under projection, be it perspective or parallel. Other characteristics which do occur in most available lists of reliable NAPs are collinearity, curvilinearity, cotermination, and parallelism. SS, which is studied more closely in this paper, is present on some lists but not on others. For example, Stevens (1980) and Kanade (1981; Kanade & Kender, 1983) investigated this property explicitly, whereas Lowe (1985) did not mention it, and Biederman (1987) added a question mark next to it.

However, mathematically showing that a property is invariant or inversely invariant is only part of the story. If NAPs are to function as a basic source of information in solving the recovery problem, they must be detected and used as such. Computer vision research has shown that this is possible in principle as well as in some working algorithms. For example, SS in the image, which results from orthographic projection of BS on a plane oriented arbitrarily in space, can be detected on the basis of the fact that the virtual lines connecting symmetrically positioned elements are still parallel under orthographic projection (parallelism is an affine invariant grouping). Moreover, the angle between these lines and the axis of symmetry in the image can be used to infer the 3-D orientation of the plane in the world. More specifically, this angle provides some constraints on the possible range for both the slant and tilt components of surface orientation relative to the viewer (see Friedberg, 1986; Hakalahti, 1983; Kanade, 1981; Kanade & Kender, 1983; Stevens, 1980).

PSYCHOLOGICAL EVIDENCE FOR DETECTION AND USE OF NAPS

With respect to human perception, not much experimental research has been devoted to discovering if NAPs are detected and used. Nevertheless, Biederman's (1987) Recognition-by-Components theory relies heavily on the requirement that NAPs are processed preattentively, because they are assumed to provide the necessary information to determine the identity of "geons" or geometric ions (primitives), supposed to be the basic building blocks of object recognition. For example (see Figure 2), a brick can be distinguished from a cylinder because it has three instead of two parallel edges, only straight edges, no curved edges, three locations where three lines coterminate, etc. From that perspective, he listed some of the "psychological evidence for the rapid use of nonaccidental relations" (pp. 119-120), but a closer look at this "evidence" reveals several problems with it.

First, it is often mere demonstration. For example, the Ames' room is seen as being derived from an assumption of symmetry that includes parallelism. Secondly, evidence of a more experimental nature is almost exclusively devoted to the detection of the regularities as such, not to their use. For example, Garner's (1974) and Pomerantz' (1977) work are cited as showing that the degree of symmetry is a readily available perceptual distinction. Biederman needs this to enable a rapid discrimination between square- and circle-shaped cross sections on the one hand, and rectangular and elliptic on the other. However, the latter kind of evidence is not directly relevant because it is focused on perfect symmetry only (which is only in particular higher-order types of symmetry preserved under projection).

Moreover, Biederman refers to experiments with visual search tasks showing that when a target differs from distractors in a NAP, the detection of that target is facilitated compared to conditions where targets and backgrounds do not differ in such properties. However, his example, curved versus straight line segments, is more exception than rule. Whereas NAPs typically imply relative positions of features in the image (e.g., parallelism, cotermination, collinearity), Treisman's original feature-integration theory (e.g., Treisman & Gelade, 1980) assumed that only targets defined by primitive feature attributes (such as orientation) could be detected preattentively and that targets defined by conjunctions of attributes required focused attention.

Forced by recently obtained data (e.g., 3-D "pop-out", Enns, 1990, 1992; Enns & Rensink, 1990a and b), later versions of her theory (e.g., Treisman, 1988; Treisman & Sato, 1990) allow fast search dependent on relative spatial positions, but only in specific cases where "emergent features" (Treisman & Patterson, 1984) arise (e.g., a triangular positioning of three line segments adds closure). However, most visual search experiments directly relevant to the issue of NAPs derived from relative spatial positions yielded slow search rates counterindicative of preattentive processing (e.g., parallel vs. converging: Treisman, 1985; circles vs. ellipses: Treisman & Gormican, 1988). There is some evidence about fast processing of spatial relations such as gaps and offsets (Foster & Ferraro, 1989), but these are not relevant to distinguish geons. Despite this foremost negative evidence, recent research with different experimental paradigms has shown that human perceivers are able to pick-up and use affine information in some situations (e.g., Cutting, 1987; 1988; Todd & Bressan, 1990), but the relevance of these experiments to the status of SS as NAP is not immediately obvious.

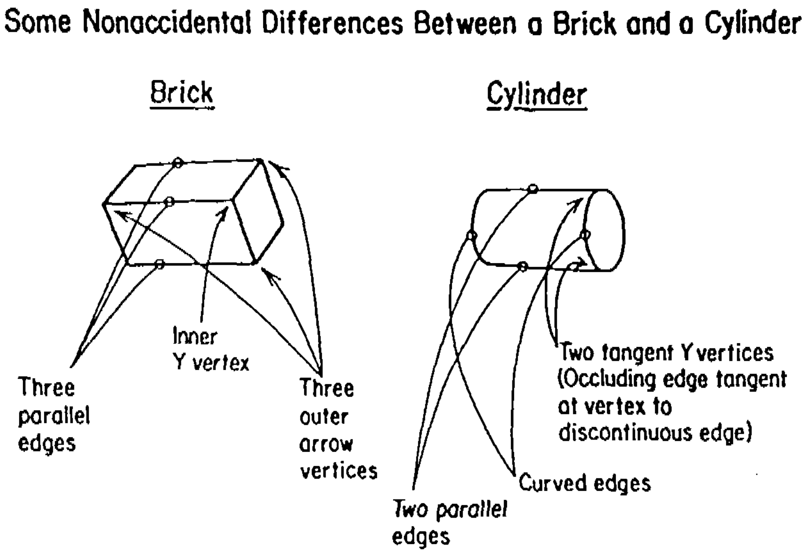

Fig. 2 NAPs are used in Biederman's (1987) Recognition-by-Components theory to distinguish geons such as a brick and a cylinder. From "Recognition-by-Components: A theory of human image understanding" by I. Biederman, *Psychological Review*, *94*, p. 121. Copyright 1987 by the American Psychological Association.

Finally, critical evidence is lacking for the use of these properties *as* NAPs. This means, given that there is a NAP in the image, is it used to infer the corresponding structure in the world? Almost no experiments directly focused on this question have been performed. Only casual observations are made, but often contradictory. For example, with respect to SS, Palmer (1983) claimed that, in cases such as an ellipse or a parallelogram, it is often readily perceived as arising from a tilted symmetric object or surface (i.e., tilted circle or top surface of a brick), whereas Attneave (1982) demonstrated a case (a dot pattern) where SS is not seen as such.

The experiments summarized in this paper are specifically devoted to finding out whether NAPs such as SS are used and, if so, by what kind of mechanism. The motivation behind it is the following. From basic mathematical work, it is known that SS is a NAP, which means that SS in the image results from orthographic projection of BS on a plane oriented arbitrarily in space, and can, therefore, be used to recover BS from it. From computer vision work, it is known that SS can be detected sufficiently reliably to be used as NAP. In addition, working algorithms exist that recover from SS the original BS as well as the slant-tilt of the plane on which it is present. From our previous experiments published elsewhere, we know that SS can be detected, but serious doubts must be formulated with respect to the speed and efficiency with which this can be achieved (see Wagemans, Van Gool, & d'Ydewalle, 1990, 1991, 1992).

The total set of eight experiments to be discussed below, can be divided in two subsets of four that differ in the kind of perceptual use that is made of SS. In the first set, the experimental task could profit from the use of SS as a NAP suggesting BS in depth. These four experiments are only reviewed here, because they will be published elsewhere (Wagemans, 1992). In the second set of four experiments, which are described here for the first time, an attempt is made to test whether subjects can infer the 3-D orientation from SS. In a sense, the latter experimental task can be regarded as a shape-from-skewed-symmetry problem, analogous to other shape-from methods such as stereo, motion, shading, and texture (for a review of the computational and psychological research of these so-called "modules", see Wagemans, 1988, 1990). The experiments about "SS as NAP" are summarized in Part II, whereas those focused on "Shape from SS" are described in Part III.

Experiments about the Perceptual Use of SS as NAP

In an attempt to remove the question mark next to SS in Biederman's (1987) list of potentially helpful NAPs, four experiments were designed in which the subjects' task was to determine whether two patterns presented side by side on a screen were the same or not, regardless of possible affine transformations of the original patterns. In other words, subjects had to assume that the patterns were present on a plane, the 3-D orientation of which could be different for the two patterns to be compared. Half of the patterns were symmetric, half random. In these conditions, BS patterns are transformed to SS patterns, and SS, if used as a NAP, might facilitate the comparison. If a SS in the image is interpreted as a BS in the world, the comparison task should be easier because there is a long research tradition showing that Gestalts, such as those afforded by BS, are easier to compare than "non- Gestalts" such as random patterns (e.g., Bagnara, Boles, Simion, & Umilta, 1983; Fox, 1975).

The experimental stimuli can be divided into eight categories. In addition to the random patterns, seven types of symmetry were used, defined by the number and orientation of the axes: four single (vertical, V; horizontal, H; left, L, i.e., rotated counterclockwise from V by 45°; and right, R, i.e., rotated clockwise from V by 45°), two double (VH or LR), and one fourfold (VHLR). In all four experiments, the affine transformation used was generated by a combination of rotation (in the plane) and compression (in the plane, which results from slant in depth). These factors were manipulated orthogonally with eight levels of orientation in the plane (0°, ±22.5°, ±45°, ±67.5°, and 90°) and four levels of compression (1, .9397, .766, .5, which correspond to 0°, 20°, 40°, and 60° slant in depth), respectively.

The four experiments differed in the way the patterns were presented. In one experiment, the patterns consisted of 24 dots located randomly or in symmetric positions in a circular area. In another experiments, the dots were connected by straight lines such that random or symmetric polygons were obtained. In two other experiments, the same patterns (dots versus polygons, resp.) were surrounded by a foursided frame which suggested the slant of the plane (e.g., a square corresponds to orthogonal viewing, a parallelogram corresponds to the affine transformed patterns suggesting oblique viewing).

The main purpose of the four experiments was to investigate if SS is helpful in a same-different comparison task. The answer was definitively positive. No matter if it was tested with dot patterns or polygons, surrounded by a frame or not, it was always easier to match a pattern with affine transformed versions of it, when the pattern contained symmetry: Less time was needed and fewer errors were made. A more detailed description of the experimental data can be found elsewhere (Wagemans, 1992).

This result supports the view that SS can be detected and used in tasks such as these, because the affine transformations that were introduced simulated orthographic projections of planes oriented in depth so that the symmetry in the dot patterns or polygons present on these planes was most of the time skewed. However, this symmetry advantage does not necessarily imply that human perceivers use SS as a NAP in doing this task. In order for this to be the case, an additional condition must be satisfied in the sense that the SS in the image had to be interpreted as being originated by a BS oriented in depth. In general, this seems to be an essential prerequisite for a particular property in the image to be useful in recovering scene properties from it.

Subjects did not have the impression of seeing the dot patterns as being slanted-tilted. In most cases, the dot patterns were perceived as being transformed in the image plane (see Figure 3). This was especially so for random dot patterns where local groupings were used as the essential information to rely on in deciding on the affine equivalence of the patterns to be compared (e.g., clusters or curvilinearity formed by three or four dots). The absence of reliable rotation-compression interactions can be interpreted as quantitative evidence for this qualitative observation. The underlying rationale is the following: If SS would be seen as BS in depth, some rotation-compression combinations should yield different effects than others. More specifically, the effects caused by pure compression (i.e., no rotation) and by pure rotation (i.e., no compression) in which one does not leave the image plane (i.e., parallel 3-D orientations of object and picture plane), should be different from those in which a slant-tilt in depth is suggested by the transformation in the image plane.

A possible reason for this negative result with respect to SS as a NAP might be the fact that dots do not necessarily force perception of coplanarity. Therefore, two ways to remediate against this were attempted. One was to make the spatial order of the dots explicit by connecting them by straight line segments (i.e., using polygons instead of dot patterns). This manipulation resulted in the expected finding of statistically significant rotation-compression

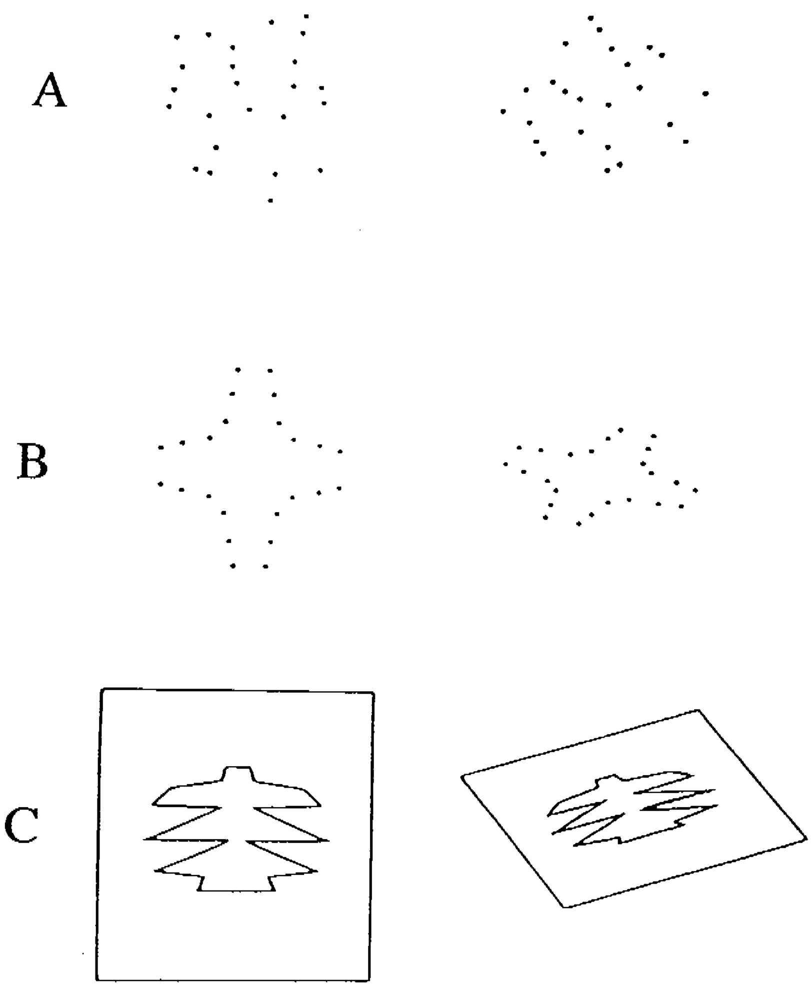

Fig. 3 Examples of the experimental stimuli used in the "SS as NAP"-study published elsewhere (Wagemans and Vanden Bossche, 1991). (A-B) Dot pattern pairs. The affine equivalence between the random dot patterns in A can be established on the basis of local groupings such as the curvilinearity. The SS in B is perceived as BS-in-depth because of the higher-order symmetry. (C) A pair of framed polygons which allows to infer the 3-D orientation.

interactions, so that SS was generally useful as well as a reliable source to infer BS oriented in depth. The second manner to improve the perceptual impression of planarity was to incorporate the dot patterns in a frame that was transformed accordingly, suggesting the contours of a planar section (i.e., a sheet of paper). This manipulation was unsuccessful. Although symmetric dot patterns were again much easier than random ones, the critical rotation-compression interaction remained absent.

Two tentative conclusions can be formulated on the basis of the findings obtained in this set of experiments. First, although symmetry is generally very helpful in comparing affine transformed patterns, SS does not really function as a NAP for dot patterns. SS polygons, and perhaps some specific cases of dot patterns surrounded by a frame, seem to be perceived as BS ones oriented in depth. Secondly, connectivity is probably the most important factor in determining this difference between dot patterns and polygons. Some observations suggest that the absence of a "Gestalt-like" form of the dot patterns forces the subjects to rely on local groupings which are qualitatively affine invariant, such as clusters based on proximity or curvilinearity. This is especially so for random dot patterns.

Experiments about the Perceptual Use of Shape from SS

In computer vision, SS is used as a source of information to recover the slant-tilt of planar surfaces oriented in depth (e.g., Friedberg, 1986; Hakalahti, 1983; Kanade, 1981; Kanade & Kender, 1983; Stevens, 1980). The experiments summarized above raise some doubts as to the psychological plausibility of this kind of shape-from method. Although symmetric patterns were systematically easier to compare with affine transformed variants, only highly regular polygons and some specific types of framed dot patterns containing SS seemed to suggest BS on a plane oriented in depth. Nevertheless, it was considered useful to investigate this more explicit use of SS, not only as a NAP, but as a cue to infer the slant-tilt of the underlying surface, more systematically. Are human perceivers able to use SS as a source of information from which to recover surface-orientation? Before trying to answer this question experimentally, a brief overview of some other psychological research concerning surface-orientation recovery will be given with two aims in mind: First, to have a somewhat broader perspective on the issue, and, secondly, to have an idea about what to expect as an outcome of these experiments.

PREVIOUS RESEARCH ABOUT SURFACE-ORIENTATION RECOVERY

Most directly focused on this topic is Epstein's work on the perception of shape-at-a-slant. He proposed a processing model, which is in fact very closely related to the old shape-slant invariance hypothesis as originally formulated by Koffka (1935). This hypothesis was suggested as an explanation of shape constancy, that is, the tendency to perceive the true shape of an object, independently of its orientation and distance relative to the viewer. Under the invariance hypothesis, the perceived true shape of an object is uniquely determined by the values of the projected shape on the retina and the perceived orientation or slant of the figure. The way this can be achieved was supposed to be by a tendency towards "Prægnanz" in the early days of Gestalt psychology, and by unconscious inference in more cognitive versions of the hypothesis (e.g., Epstein, 1973; Rock, 1977).

In Epstein's model (Epstein & Lovitts, 1985), the perceived shape-at-a-slant (i.e., the output) is computed by a shape-slant algorithm working on the basis of two representations, one of the projective shape and one of the orientation information, which are each encoded on the basis of specific registered variables, generated by the distal shape-at-a-slant (i.e., the input). Most of the subsequent research by Epstein and his associates was devoted to finding out until what point in this information-flow diagram the processing is automatic and from what point it demands attention. Initial results of this work suggested partial automaticity, that is, automatic encoding of the essential information, whereas attention would be required for their combination (Epstein & Babler, 1989; Epstein & Broota, 1986; Epstein, Hatfield, & Muise, 1977; Epstein & Lovitts, 1985). However, a recent series of experiments using the visual search paradigm, which typically allows to discriminate preattentive from attentive types of processing, yielded somewhat ambiguous results (Epstein & Babler, 1990).

Although the experiments reported here are not directly devoted to the question as to where attention starts playing a role, the underlying motivation as posed above can be reformulated to fit into this scheme. If the projected shape is skewed symmetric and the perceived orientation could be determined on the basis of this (as implied by current computer vision algorithms), then it should be easier to arrive at the correctly perceived distal shape (i.e., BS on a slanted-tilted plane). An alternative formulation, implying a reverse kind of influence, could be that it would be easier to derive the perceived orientation if the resulting perceived shape could be more regular than in other cases (i.e., BS instead of random). This alternative route, which leaves the strictly sequential path as implied by the more cognitive versions stressing the role of unconscious inferences, is very Gestalt-like in flavour. In fact, the suggestion is that a tendency towards a stronger Gestalt ("Prægnanz") drives the surface-orientation recovery. Palmer (1983) formulated a theory of perceptual organization which explicitly enables this kind of influences via the selection of a reference frame based on maximizing symmetries.

Regardless of the direction of influences, the basic question is whether evidence exists to suggest that symmetry can play some role in this process of recovering surface orientation and distal shape from 2-D input images. First, at a theoretical level, Shepard (1981) indicates the possibility:

> "The mapping from three dimensions down into two is necessarily singular and so does not possess a unique inverse. The rules of formation must therefore incorporate some principles for selecting one out of the infinite set of inverse transformations that will 'retroject' any given 2-D projection back out into the 3-D world.

> The perceptual system may select a 3-D interpretation that optimizes simplicity, regularity, or symmetry" (p. 307).

Along with this theoretically inspired suspicion, Shepard (1981) mentions an experiment in which the alternate presentation of a random polygon with a 50% compressed version of it was giving rise to the perceptual impression of a rigid (presumably 60°) rotation in depth. Although the apparent motion paradigm might have been critical to obtain this result, similar compressions will be used here to attempt to suggest slant in depth.

A second example of supporting statements or evidence is Rock's explanation of the difference between his own findings suggesting egocentric coding and Marr's (1982) 3-D object model implying viewer-independent representation:

> "Processing always begins on the basis of the retinally based egocentric descriptions. If, however, there is information to the effect that the description does not do justice to the object, or afford the best description of it, further processing occurs, the aim of which is to achieve a more object- or environment-centered description. In the case of a 2-D object seen at a slant, awareness of its slant informs us that a better description would be based on a view that was normal to its surface (...) If the shape is not too complex, it is not too difficult to achieve that description" (Rock & DiVita, 1987, p. 292).

Perhaps SS patterns obey this condition. To perceive BS at a slant is easier than the percept of SS as it is.

More specific evidence in favour of this hypothesis was obtained by Biederman and his colleagues some years ago (King, Meyer, Tangney, & Biederman, 1976). The starting point of their research was the observation that shape constancy is less successful at larger angles of rotation (i.e., tendency towards underestimation of slant). In an attempt to investigate this effect experimentally, a strange phenomenon was observed: When symmetric figures (circles and squares) were rotated through larger angles, the perceived shape was less symmetric due to a tendency towards the projective shape. If, however, less symmetric figures (ellipses and rectangles) were rotated through larger angles, the perceived shape was more symmetric, which is clearly in contrast with the projective shape. Subsequent experiments showed that this is not a response bias or a strategic effect but a genuine perceptual tendency towards higher symmetry.

More recently, Stevens (1980, 1983) asked his subjects to judge the orientation of a surface by positioning a needle to indicate the surface normal (i.e., pointing out the surface perpendicularly). The stimuli used were simple *oblique crosses and parallelograms*. The subjects' tilt estimates corresponded closely to an interpretation in depth in which the obtuse intersection of two line segments is seen as a foreshortened right-angle intersection in 3-D and in which the line segments are seen as being of equal length in 3-D. In short, in tilt judgements as well as in slant estimates (e.g., Attneave, 1972; Attneave & Frost, 1969), there was a tendency to regularize the perceived shape towards equal lengths and right angles.

In summary, some theoretical positions as well as both kinds of experimental evidence discussed above, suggest that a tendency towards higher "Prægnanz" might enable SS to be seen as BS in depth and to be used as a reliable source of information for surface-orientation recovery. However, the stimuli used in these experiments were very simple, whereas the experiments in the previous paragraph suggested that the usefulness of SS might be seriously restricted to this particular kind of simple stimuli. It remains to be seen, therefore, if similar results would also be obtained with more complex patterns such as dot patterns or polygons of higher complexity than the parallelograms used before.

In addition to the latter distinction between dot patterns and polygons, a second dichotomy underlies the manipulations in this set of four experiments. Older work on slant estimation showed that absolute slant estimates are rather inaccurate, whereas relative judgements are quite good (see Cutting & Millard, 1984, for a review). In both paradigms used here, the slant suggested by the affine transformed patterns had to be judged relative to the slant suggested by a surrounding frame. Experiments 1 and 2 tested performance rather indirectly with same-different judgements, whereas Experiments 3 and 4 required the subjects to actively adjust the slant of the frame to the one suggested by the affine transformed patterns.

Experiment 1

METHOD

Subjects

Two highly practiced and two naive subjects participated in this experiment. The two practiced subjects were randomly selected from a pool of four consisting of the author, one graduate and two undergraduate students involved in this research project. The two naive subjects were volunteers from a pool of undergraduate students having to participate at least in four short experiments or one 4-h experiment in fulfilment of a course requirement. All observers had normal or corrected-to-normal vision.

Stimuli

In each trial two dot patterns, each consisting of 24 dots in a circular area with a 5.5 cm cross section, were presented on the screen, one in the left half, the other in the right half, separated by a centre-to-centre distance of 10 cm. The one in the left half of the screen was always untransformed (i.e., the original pattern), the right one was an affine transformed version of it. How they were created is described in detail in Wagemans et al. (1991). For the purpose of this experiment, it is sufficient to know that half of the original dot patterns were completely random, whereas the other half were symmetric. Seven types of symmetry were used: In single symmetry, twelve randomly positioned dots were reflected about an axis oriented V, H, L, or R. Two kinds of double symmetry (VH or LR) were created by first reflecting six randomly positioned dots about a V or L axis, and then reflecting the resulting 12 dots pattern about the orthogonal axis (H or R). In quadruple symmetry, a pseudorandom collection of three dots was reflected about all four axes (VHLR).

In all four experiments, the affine transformation used was generated by a combination of rotation (in the plane) and compression (in the plane, which results from slant in depth). These factors were manipulated orthogonally with eight levels of orientation in the plane (0°, ±22.5°, ±45°, ±67.5°, and 90°) and four levels of compression (1, .9397, .766, .5, which correspond to 0°, 20°, 40°, and 60° slant in depth), respectively.

The dot patterns were surrounded by a frame suggesting the slant of the plane on which the pattern was present. The untransformed patterns at the left side of the screen were included in a 8 cm sided square, whereas the transformed patterns were surrounded by a rectangle or a parallelogram, the orientation and size of which was determined by the rotation angle and the compression factor of the pattern. In contrast with the previous study, however, the transformation of the frame was in accordance with the affine transformation of the pattern in only half of the trials. In the other half, both parameters (rotation angle and compression factor) differed by one level. For example, when the pattern was rotated over a 22.5° angle and compressed by a factor 0.766, the frame could be rotated over a 0° or 45° angle and compressed by a factor 0.9397 or 0.5.

These one-level differences were chosen for three reasons. First, pilot work had determined that this kind of difference was sufficiently large to be easily discriminable (except maybe for the difference between the smallest compression and no compression, but this was always combined with a one-level difference in rotation angle). Secondly, by restricting the set of possible differences to such small deviations, problems were avoided that could be associated with invariances of a square under 90° rotations. Finally, such small differences made it highly unlikely to create frames which would be overlapping with the patterns themselves. Despite this caution, a possible cue to the difference of pattern and frame could have been an extremely small distance between the peripheral dots or lines in a pattern and the frame. If used, this kind of artifact will show up in the data.

For each of the 32 rotation-compression combinations used to transform the dot patterns, one different rotation-compression combination was selected randomly from the set of all possible one-level differences (the number of

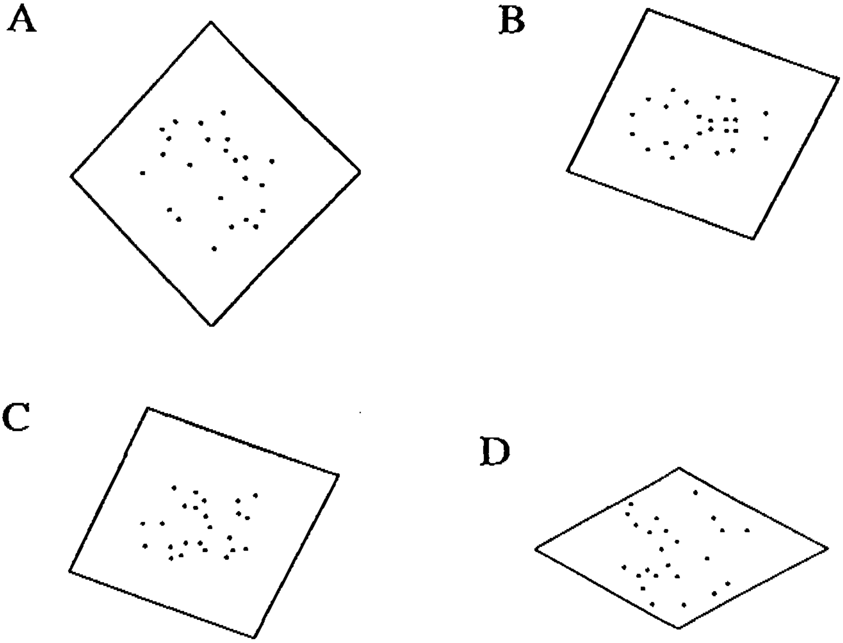

Fig. 4 Examples of stimuli in Experiment 1. (A) Symmetric dot pattern with corresponding frame. (B) Symmetric pattern with wrong frame. (C) Random dot pattern with corresponding frame. (D) Random pattern with wrong frame. In the latter case, a Different response can be given fast because of the proximity cue.

which varied from one to four). Each of these 32 possibilities occurred three times with the corresponding frame and three times with an incorrect frame for each type of patterns (seven symmetric and one random). Because random dot patterns were made to occur as frequently as symmetric ones, the total number of trials was 2,688. In each trial, a different pair of dot patterns was used (unless the response for the trial was wrong, cf. infra). Examples of trials are given in Figure 4.

Apparatus

The coordinates of the original dot patterns were generated by a C program on a SUN-3 Workstation with a Motorola MC 68881 floating point board (as described in Wagemans et al., 1991). The coordinates of all patterns were stored in separate files that could be accessed by a Turbo Pascal program automating the experiment on an IBM-AT-compatible with an Intel 80286 processor and a Hercules graphics board. The dot patterns were presented as black dots on a grey background on a monochrome screen (with a fast white phosphor, used in reverse mode) with a 60 Hz temporal resolution and a 720 x 348 spatial resolution. The subjects were seated at a distance of 114 cm on a chair with adaptable height to get their eyes in front of the middle of the screen. At that viewing distance the individual dots, the patterns constituted by them, and the whole stimulus area (both patterns plus the separating space) subtended 6.6 arc min, 2.75°, and 7.75° of visual angle, respectively. The room was completely dark except for the light reflecting from the screen. A headrest was used to prevent head rotations.

A trial was constituted from the following sequence of events. First, a black fixation cross was presented during 750 ms in the centre of a grey background. Secondly, after a 500 ms blank interval, two dot patterns surrounded by a frame were presented: The untransformed one at the left side of the fixation cross, the transformed one at the right side of it, either with a corresponding frame (the same rotation-compression combination as the pattern) or a slightly different one. Everything (dot patterns and frames) remained on the screen until the subject pressed one of two response buttons on a panel connected with the PC configuration. RTs were recorded from stimulus onset until button press to the nearest 0.1 ms. Each answer was instantaneously evaluated by the computer so that immediate feedback could be given. A correct answer was followed by a 500 ms high-frequency tone (660 Hz), a false response by a 500 ms low-frequency tone (440 Hz). The only reason why this feedback was provided was to keep the motivation and arousal of the subjects at an optimal level. Secondary learning effects as a result of this feedback procedure were averaged out by randomizing block and trial orders.

Task and procedure

The subjects' task was to indicate whether the slant of the plane as suggested by the square, rectangle or parallelogram corresponded with the slant of the dot pattern as suggested by the affine transformation that was applied to the right dot pattern. This could be inferred by comparing the transformed version of the dot pattern or with the untransformed one at the left side of the screen. Subjects responded by pressing one of two buttons on a response panel, the identity of which could be chosen by the subjects (i.e., left one indicating *Same* responses, the right one *Different*, or the reverse). Although the verbal instructions sounded rather complex to the subjects, the task was quite clear after a block of 168 practice trials which were similarly constituted as the experimental ones.

The total number of trials (i.e., 2,688) was divided in 16 blocks of 168 trials that could be run without a break. Trials with wrong responses were retaken at a random position in the remainder of the block. This was done to maximize the chance that the numbers of RTs associated with a correct response were equal in all conditions. However, to prevent that a block of trials would become too long as a result of this aspect of the procedure, the number of repetitions was limited: If the trial was not answered correctly after three attempts, it was not repeated further. The RT used in those cases was the

TABLE 1
ANOVA on RTs for Experiment 1

Effect	*df*	*F*	*p*	%
Kind	1,3	20.71	.05	11.86
Pattern	6,18	18.16	.00001	4.19
Rotation	7,21	24.37	.00001	19.28
Re × K	1,3	38.10	.01	1.11
Re × Ro	7,21	7.43	.0005	2.56
K × P	6,18	12.62	.00001	2.96
K × Ro	7,21	17.30	.00001	9.58
P × Ro	42,126	1.54	.05	1.80
Ro × C	21,63	2.56	.005	3.86
Re × K × Ro	7,21	4.61	.005	1.40
Re × Ro × C	21,63	3.66	.00001	3.16
K × P × Ro	42,126	2.69	.00001	3.25
K × P × C	18,54	1.95	.05	0.97
K × Ro × C	21,63	3.05	.0005	1.99
P × Ro × C	126,378	1.26	.05	3.94
Re × K × P × Ro	42,126	1.79	.01	1.60
Re × P × Ro × C	126,378	1.29	.05	4.25
K × P × Ro × C	126,378	1.33	.05	4.49

one associated with the final incorrect attempt. Each block, taking about 15 or 20 min, was a random sample from the total number of trials with the constraint that it contained an equal number of positive and negative trials, both subdivided in equal numbers of all kinds of symmetric and random patterns. All subjects took three or four blocks in one 1-h session with only minor (e.g., 2 or 3 min) breaks between the blocks. Sessions were distributed across several days within a one week period.

RESULTS

The data were analyzed by an ANOVA on RTs associated with a correct responses in a within-subjects design with five orthogonal factors: Response (2 levels: same vs. different), Kind of patterns (2 levels: symmetric vs. random), Pattern type (7 levels: VHLR, VH, LR, V, H, L, R), Rotation (8 levels: 90°, 67.5°, 45°, 22.5°, 0°, -22.5°, -45°, and -67.5°), and Compression (4 levels: 1, .9397, .766, and .5). Numbers of errors made in each condition were also analyzed. The results of this analysis will not be reported because they were very similar to the RTs. In none of the experiments reported here evidence for speed/accuracy tradeoffs was obtained. Of all 2,668 observations per subject, only a small number was not associated with a correct response but with the last incorrect one. These numbers were 21, 48, 14, and 23 for the four individual subjects. Means based on excluding these "incorrect" RTs were

TABLE 2
Effects of Pattern Type on Response Times (in s) in all four Experiments

	Experiment			
	1	2	3	4
VHLR	1.4	0.8	5.0	3.6
VH	1.4	0.8	6.0	5.0
LR	2.0	1.1	6.9	6.0
V	1.8	0.8	5.4	4.1
H	2.0	0.9	6.3	5.3
L	2.3	1.1	6.8	5.4
R	2.3	1.2	6.9	5.5
Random	3.9	1.4	10.1	5.6

very similar. Pattern type does not make much sense for the random patterns. Nevertheless, it was incorporated as such to obtain orthogonally manipulated factors. An additional argument might be that the way the patterns were generated (together with the symmetric ones for each type of symmetry) could have caused particular "features" to have "sneaked" in each set of random patterns. However, sometimes results averaged across this Pattern type variable will be reported for the random patterns in order not to overload the general picture with too much superfluous detail. For each main and interaction effect, the interaction with subject was taken as an error term. Proportion of explained variance will be given as an indication of the size of an effect. All significant effects are indicated in Table 1. In the following, some more specific results based on a posteriori comparisons will be summarized. Although the statistical details of all these effects will not be given, all discussed effects are significant at the p <.05 level.

Whereas there was no main effect of *Response* (F < 1), the main effect of *Kind*, explaining 12% of the variance, was as expected: Symmetric patterns yielded much shorter RTs than random ones (1892 vs. 3921 ms). The interaction with *Pattern* is due to the fact that type of pattern had no effect on the random patterns (cf. supra). With respect to the different types of symmetry, there was a large effect of number of axes: Quadruple symmetries yielded much shorter RTs than double or single ones (1373, 1536, and 2005 ms, resp.), which also differed significantly from each other. There was also an effect of axis orientation (see Table 2, first column): VH differed from LR, and single symmetry was easier about a V axis than about a L or R axis. The two-way interaction between Response and Kind is the consequence of the fact that it was easier to respond *Same* for symmetric patterns and *Different* for random ones. This effect did not interact with Pattern (p > .10).

The main effect of *Rotation* was the most important effect in terms of the percent explained variance (20%). For both the symmetric and the random patterns, this implied longer RTs for angles deviating more from V. In

TABLE 3
Effects of Rotation on Response Times (in s) for Different Kinds of Patterns in all four Experiments

	Experiment							
	1		2		3		4	
Rotation	SYM	RND	SYM	RND	SYM	RND	SYM	RND
90°	1.9	4.7	0.9	1.6	7.3	13.3	6.1	5.8
67.5°	1.9	4.4	1.0	1.4	7.4	12.9	5.8	7.2
45°	1.9	4.2	1.0	1.4	6.1	10.1	5.1	6.0
22.5°	1.9	3.2	1.0	1.3	5.4	8.1	4.6	5.0
0°	1.6	2.6	0.9	1.0	3.6	4.8	3.0	2.5
-22.5°	1.8	3.4	1.0	1.3	5.5	8.3	4.7	4.6
-45°	2.0	4.4	1.0	1.4	6.8	11.3	5.2	6.1
-67.5°	2.1	4.5	1.0	1.5	7.5	12.4	5.6	7.7

addition, the rotation effect was much more pronounced for the random patterns than for the symmetric ones (see Table 3, first two columns). For both kinds of patterns, angles close to 90° produced somewhat smaller RTs than would be expected on the basis of a linear effect from deviation from V. This was apparent from significant quadratic and quartic components in the trends, although these were only just reliable for the symmetric patterns ($p < .05$), whereas they were both highly reliable ($p < .00001$) for the random ones.

Furthermore, this interaction was modified by Response. For the symmetric patterns, the quadratic and quartic Rotation effects were present only for the *Same* responses, not for the *Different* ones, whereas both components remained reliable for both response types for the random patterns. A most peculiar finding is that there was a significant decrease in RTs at both 67.5° rotations for the *Different* random patterns and not for the *Same* responses, nor for the symmetric ones. As suggested above, this most probably resulted from an artifact created by the procedure of rotating and compressing the frame by a one-level difference which caused some dots to be positioned very closely to the frame (see Figure 4). The same will be noted for some higher-order interactions.

The effect of Rotation was also somewhat different for the different types of symmetric patterns. For example, the quadratic trend was not reliable for the highly regular patterns (VHLR, VH, and LR), whereas the quartic component disappeared for other types (LR, V, L, and R). Interesting differences between quadruple and double symmetry appeared also in the fact that the difference between 0° and 90° rotations on the one hand, and ±22.5° and ±67.5° on the other, were reliable only for the former and not for the latter type of symmetric patterns. The reverse was true for the difference between the ±45° rotations versus the intermediate ones (±22.5° and ±67.5°).

TABLE 4
Effects of Rotation and Compression on Response Times (in ms) for Different Kinds of Patterns in Experiment 1

Symmetric/Same ($X = 1781$)

	Rotation Angle							
Compression	90°	67.5°	45°	22.5°	0°	−22.5°	−45°	−67.5°
1	1721	1805	1660	1698	1033	1644	1844	1900
.9397	1571	1831	1808	1658	1262	1537	1888	2054
.766	1777	1899	1614	1663	1569	1878	1736	1923
.5	1965	1835	2175	2079	1512	1849	2449	2148

Symmetric/Different ($X = 2003$)

	Rotation Angle							
Compression	90°	67.5°	45°	22.5°	0°	−22.5°	−45°	−67.5°
1	2350	2439	2028	2201	1841	2177	2011	2082
.9397	1915	1778	2097	1983	1575	2206	1986	2005
.766	2115	1211	1921	1484	1718	1454	2204	2040
.5	2145	2679	2018	2179	2025	1973	1978	2287

Random/Same ($X = 4006$)

	Rotation Angle							
Compression	90°	67.5°	45°	22.5°	0°	−22.5°	−45°	−67.5°
1	5025	5269	4443	3622	1958	3707	4779	5056
.9397	5106	4685	4578	3469	1807	3522	4660	4623
.766	4578	4357	4032	2962	2026	3244	4539	5068
.5	4615	4857	4020	3179	2795	3389	3770	4457

Random/Different ($X = 3835$)

	Rotation Angle							
Compression	90°	67.5°	45°	22.5°	0°	−22.5°	−45°	-67.5°
1	5162	4971	4448	3137	3219	3615	4532	5331
.9397	4955	4657	4170	3265	3159	3429	4169	5034
.66	4657	2068	4303	2215	3128	2376	4303	2293
.5	3476	4027	4028	3823	2957	3809	4244	3778

Although the main effect of *Compression* was not statistically significant ($p > .25$), it was involved in some reliable higher-order interactions with Response, Kind, Type, and/or Rotation (see Table 4). Because only small percentages of explained variance (i.e., less than 5%) were associated with these effects, only the most salient differences are described. The Rotation-Compression interaction effect was completely different for the different Response- Kind combinations. Whereas almost no pairwise differences or trends were present for the symmetric kind of patterns (especially for the *Same* responses), large effects were obtained for the random patterns. More specifically, *Same* random patterns yielded large quadratic and quartic rotation effects, although their relative weight could fluctuate somewhat as a function of the compression level. For the *Different* random patterns, the quadratic components disappeared for the large compression factors (i.e., .766 and .5), whereas the quartic components disappeared for the small compression factors (i.e., 1 and .9397). Furthermore, the differences between ±67.5° versus ±45° and 90° were only reliable for a compression of .766, as well as those between ±22.5° versus 0°. Again, the proximity artifact is the most likely explanation. Straightforward calculations showed that this proximity cue occurred most often for these specific rotation-compression combinations.

DISCUSSION

As in the experiments summarized in Part II, there was a large difference between symmetric and random dot patterns with respect to the time that was needed to decide whether their orientation in depth corresponded with the one suggested by the frame. Moreover, previously observed differences due to number and orientation of axes were also replicated. As such, these findings seem to argue for the advantage offered by symmetry in performing this task. Additional findings, specific for this experiment, such as the *Same* advantage for symmetric patterns and the *Different* advantage for random patterns, fit into this general scheme. The differential effects of rotation angle for quadruple symmetry and double VH symmetry also seem to argue for the perceptual processing and use of symmetry.

However, there are reasons to believe that subjects are not really recovering surface orientation from SS. First, the way the task is conceived makes it possible for the subjects to remain within the image plane. Especially in the cases where symmetric patterns are surrounded by a frame, parallelism between one of the frame's sides and the virtual line connecting both most peripheral symmetrically positioned dots is sufficient to decide *Same*, without having to recover the surface orientation in depth (see Figure 4). Furthermore, this explains the difference between V and H versus L and R symmetries. This alignment is not possible in the latter cases. In addition, subjects did not have the impression of seeing BS on a tilted plane, except in the highly regular cases (quadruple or double symmetry).

As suggested before, the patterns were probably too complex to afford a salient global structure (Gestalt) which could, as a whole, be perceived as being oriented in 3-D space. In the previous experiments (Wagemans, 1992), the introduction of a frame around the dot patterns did not facilitate this kind of planar Gestalt formation very much. Moreover, in contrast with the previous experiments, this manipulation enabled artificial cues to rely on in performing the task. For some types of symmetric patterns, a simple alignment could suffice (note that this requires the symmetry to be detected and used), whereas the distance between the most peripheral dots and the frame borders could be used to decide *Different* for random patterns.

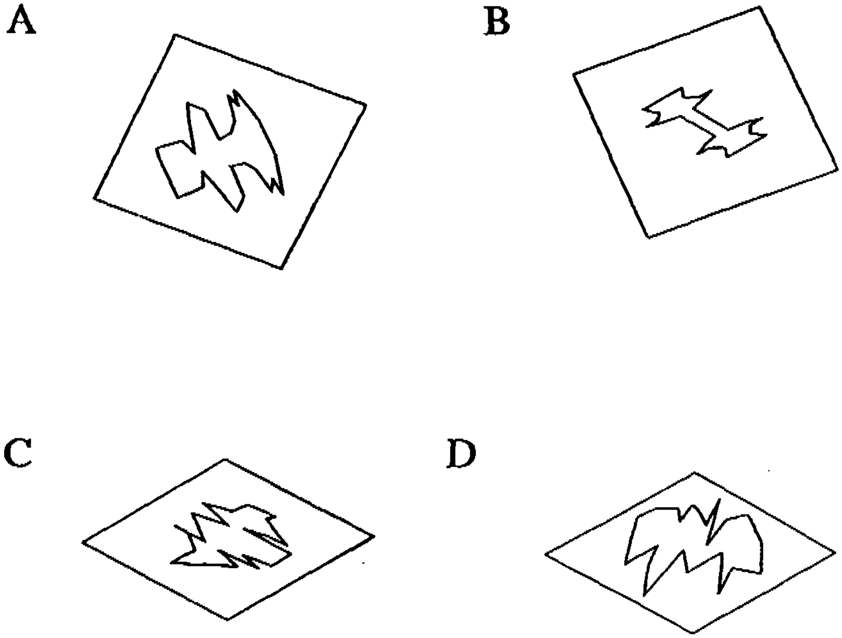

Fig. 5 Examples of stimuli in Experiment 2. (A) Symmetric polygon with corresponding frame. (B) Symmetric polygon with wrong frame. (C) Random polygon with corresponding frame. (D) Random polygon with wrong frame. In the latter case, a Different response can be given fast because of the proximity cue.

In short, evidence for or against the use of SS to recover surface orientation is hard to provide on the basis of these unreliable data. The suspicion is that a recovery process only automatically "triggers off" for highly regular patterns, whereas simple perceptual "tricks" such as the qualitative measurement of distance or parallelism in the image plane are used if this first stage fails. A way to corroborate this suggestion might be to enhance the Gestalt formation by connecting the dots by straight line segments. In the previous study, closure and explicit spatial order were shown to contribute to the saliency of SS as BS in depth. Therefore, in Experiment 2 polygons were used instead of dot patterns.

Experiment 2

METHOD

Stimuli

In each trial an original polygon and an affine transformed version of it were presented side by side on the screen: The untransformed one in the left half of the screen, the other one at the right. Both were surrounded by a frame: A square for the untransformed polygon and a square, a rectangle, or a parallelogram for the transformed one. The polygons were created by

TABLE 5
ANOVA in RTs for Experiment 2

Effect	*df*	*F*	*p*	%
Kind	1,3	37.69	.01	34.73
Pattern	6,18	9.73	.0001	4.41
Rotation	7,21	16.92	.00001	6.76
Re × P	6,18	7.25	.0005	1.01
Re × Ro	7,21	6.85	.0005	2.49
K × P	6,18	7.95	.0005	4.28
K × Ro	7,21	7.88	.0001	4.97
Ro × C	21,63	3.83	.00001	2.92
Re × K × P	6,18	3.00	.05	0.42
Re × P × C	18,54	1.94	.05	0.82
Re × Ro × C	21,63	3.25	.0005	1.95
K × P × Ro	42,126	1.91	.005	2.19
K × Ro × C	21,63	1.82	.05	1.81
P × Ro × C	126,378	1.27	.05	4.11
Re × P × Ro × C	126,378	1.47	.005	4.66
Re × K × P × Ro × C	126,378	1.28	.05	4.73

connecting the 24 coordinates generated for the dot patterns used in Experiment 1 in such a way as to prevent intersections and to preserve their type (e.g., random, V-symmetric, quadruply symmetric, etc.). As in Experiment 1, 96 polygons were generated for each kind of symmetry. In addition, 1,344 random polygons were made to have equal numbers of symmetric and random patterns. Again, half of the trials showed transformed polygons surrounded by a corresponding frame (with the same rotation-compression combination), whereas in the other half the slant of the plane suggested by the frame differed slightly from the one suggested by the affine transformation applied to the polygon. Examples are shown in Figure 5.

Subjects, apparatus, task and procedure

The rest of the method used in this experiment was completely analogous to that of Experiment 1, except for the fact that a block took only about 10 or 15 min, so that four or five blocks could be taken in a 1-h session. Again, sessions were spread over several days within a one week period.

RESULTS

The same kind of analyses as in Experiment 1 was performed. In Experiment 2, the numbers of "incorrect" RTs were 1, 35, 48, and 1 for the four individual subjects. These constituted only a minor fraction of the total number of trials (i.e., 2,688). Again, the means on the "correct" RTs only were quite similar. All significant ($p < .05$) effects are indicated in Table 5. Whereas there was

TABLE 6

Effects of Rotation and Compression on Response Times (in ms) for Different Kinds of Patterns in Experiment 2

Symmetric/Same ($X = 934$)

	Rotation Angle							
Compression	90°	67.5°	45°	22.5°	0°	−22.5°	−45°	−67.5°
1	790	893	979	1098	709	1078	1066	925
.9397	768	1007	882	862	751	867	976	1006
.766	873	877	877	997	940	876	941	909
.5	999	925	992	1076	980	1029	999	946

Symmetric/Different ($X = 1006$)

	Rotation Angle							
Compression	90°	67.5°	45°	22.5°	0°	−22.5°	−45°	−67.5°
1	1063	1160	963	906	1063	1016	937	951
.9397	1129	992	935	1000	1034	1010	977	1034
.766	1016	796	1111	982	924	979	1157	937
.5	*962*	*1247*	*993*	*903*	*919*	*959*	*1068*	*1088*

Random/Same ($X = 1362$)

	Rotation Angle							
Compression	90°	67.5°	45°	22.5°	0°	−22.5°	−45°	−67.5°
1	1778	1557	1622	1415	824	1274	1384	1647
.9397	1630	1507	1337	1576	774	1536	1334	1615
.766	1397	1357	1534	1253	865	1364	1455	1648
.5	1101	1368	1414	1221	881	1181	1288	1443

Random/Different ($X = 1395$)

	Rotation Angle							
Compression	90°	67.5°	45°	22.5°	0°	−22.5°	−45°	−67.5°
1	1737	1505	1455	1433	1106	1381	1634	1571
.9397	1725	1414	1458	1264	1074	1368	1566	1605
.766	2028	1130	1438	993	1302	939	1394	1114
.5	1448	1577	1210	1515	1224	1466	1130	1438

no main effect of *Response* ($F < 1$), the main effect of *Kind* explained the foremost largest proportion of systematic variance (35%). As in Experiment 1, symmetric patterns yielded much shorter RTs than random ones (970 vs. 1378 ms). Again, the interaction with Pattern is due to the fact that type of pattern had no systematic effect on the random patterns. With respect to the different types of symmetry, there was an effect of number of axes: Quadruple symmetries yielded shorter RTs than double or single ones (837, 960, and 1009 ms, resp.), but the latter did not differ significantly from each other ($p > .15$). However, this effect of number of axes was seriously modified by an effect of axis orientation (see Table 2, second column). For example, there were large differences between VH and LR and between single V and H symmetry versus L and R symmetry. As in Experiment 1, the two-way interaction between Response and Kind is the consequence of the fact that it was easier to respond *Same* for symmetric patterns and *Different* for random ones. In addition, there was a reliable third-order interaction with Pattern, but this explained less than a half percent of the variance.

The main effect of *Rotation* was less important than in Experiment 1 (only 7% of the systematic variance). In fact, only for the random patterns were large and systematic effects observed (see Table 3, next two columns). More specifically, the quadratic and quartic trends were highly reliable. Of the four trend components tested for all types of symmetry, only two were reliable at $p < .05$: The trend caused by rotation was quartic for the quadruple symmetries, and linear for the single symmetries about a R axis.

Although the main effect of *Compression* was not statistically significant ($p > .25$), it was involved in some reliable higher-order interactions with Response, Kind, Pattern, and/or Rotation (see Table 6). As before, only the most salient differences will be noted. The Rotation-Compression interaction effect was completely different for the different Response-Kind combinations. In contrast with the symmetric patterns, for which almost no reliable pairwise differences or trends were present, large effects were obtained for the random patterns. More specifically, *Same* random patterns yielded large quadratic rotation effects for all compression levels, whereas the quartic components only reached statistical reliability when compression was 1 or .766. For the *Different* random patterns, the quadratic components remained reliable only for the small compression factors (i.e., 1 and .9397), whereas the quartic components were marginally significant ($.05 < p < .10$) at all compression levels. As in Experiment 1, the proximity artifact is the most likely explanation.

DISCUSSION

In general, the RTs needed to perform the task with polygons instead of dot patterns were much shorter (from 2906 ms in Experiment 1 to 1174 ms in Experiment 2). Nevertheless, the difference between symmetric and random patterns was still obtained, as were the effects of number and orientation of the axes on the processing times for the former kind of patterns. As such, this seems to corroborate the motivation for this experiment: The use of polygons instead of dot patterns has the supposed beneficial influence.

However, as in Experiment 1, there are reasons to doubt whether subjects have really recovered surface orientation in depth to perform the experimental task. It was in some cases perfectly possible to rely on particular cues within the image plane, and some specific results seem to offer evidence for their use. For example, the higher-order interactions caused by decreased RTs for random polygons which differed from their frames at particular Rotation-Compression combinations argues for the use of proximity of peripheral straight line segments to the frame. Likewise, the absence of reliable or systematic rotation effects for the symmetric polygons suggests that recovery was not the underlying process.

In addition to the results which replicated those found in Experiment 1, some effects specific to this experiment are in line with the above reasoning. For example, the fact that the effect of orientation of the axis of symmetry was much more pronounced in Experiment 2 than in Experiment 1, might be due to the presence of real lines to align with the frame instead of virtual lines connecting two peripheral dots. The absence of a difference between VHLR, VH, and V corroborates this suspicion: If mere local parallelism in the image plane is used, it is quite natural that additional global regularity does not add anything to the efficiency of the process.

At this point, one would normally stop using this kind of task and stimulus material because it does not permit to test what it was supposed to. However, this series of four experiments was planned as such before they were run and data were not analyzed until all four experiments were finished. The original plan was to investigate two modes of recovering surface orientation: One which can be described as implicit, indirect, and passive; and one explicit, direct, and active. In Experiments 1 and 2, the first mode should have been studied by a paradigm in which subjects had to determine only whether the surface orientation as suggested by an affine transformed four-sided frame was the same as the one implied by the affine transformation of a pattern in it. This does not require that the surface orientation be expressed or measured quantitatively by specific parameter settings. Experiments 3 and 4 attempt to do this by asking the subjects to change the parameters defining the affine transformation of the frame in order to simulate the 3-D orientation of the plane of the affine transformed dot patterns or polygons.

Experiment 3

METHOD

Subjects

In addition to two subjects who were highly experienced in this and related tasks, four undergraduate students at the University of Leuven participated in the experiment. The latter were unpracticed and naive with respect to the purpose of the experiment. Because of the time consuming aspect of the procedure used in Experiments 3 and 4, two naive subjects were asked to perform half of the total number of trials that could be managed by a single subject in Experiments 1 and 2. For each naive subject, half of the blocks were selected randomly out of the 16-blocks set. In addition, one block was performed by both subjects together providing a full set of trials, to determine if they differed significantly in general performance level (average number of errors and mean response times). This was not the case, so that trials could be merged. All observers had normal or corrected-to-normal vision.

Stimuli and apparatus

The stimuli used in this experiment were the same as in Experiment 1, except that the frame around the affine transformed versions of the dot patterns was always initially a square. In all but one of the 32 rotation-compression combinations, therefore, the slant suggested by this square frame differed from the one suggested by the affine transformation applied to the dot pattern in it. In this experiment, however, the frame could be adjusted by the subject using the arrow keys of an extended keyboard.

Task and procedure

The task of the subjects was to adjust the frame around the affine transformed versions of the dot patterns to bring its suggested slant in correspondence with the slant suggested by the affine transformation that caused the difference between the dot pattern at the left side of the screen and the one at the right. Subjects could do this by pressing the arrow keys of an extended keyboard: The horizontal arrows could be used to change the rotation angle (L = counterclockwise, CCW; and R = clockwise, CW), whereas the vertical arrows could be used to alter the compression factor (the arrow pointing upwards increased the compression yielded by slanting the plane backwards, whereas the arrow pointing down did the reverse).

The step size of the arrows corresponded with the levels of the transformation parameters used to generate the affine transformed dot patterns. Likewise, the range of potential changes was limited to avoid rotation angles larger than 90° CCW and 67.5° CW and compression factors larger than 1 and smaller than 0.5. Of course, all arrows could be used as often as wanted and subjects could undo a particular change by pressing the opposite arrow key. When subjects were satisfied with the slant suggested by the frame (i.e., when judged to correspond with the slant suggested by the affine transformation applied to the dot pattern), they had to press the space bar to fix their setting and a response button on the panel connected with the PC configuration to get feedback and to initiate the next trial. Response times were measured from stimulus onset until button press.

In cases of invariance under rotation in the image plane, two principles were followed. If the pattern itself was invariant under pure rotation (i.e., 90° for VHLR), the subjects did not have to rotate the frame, of course, because there was nothing in the available information to tell them that a rotation had taken place. On the other hand, if the frame around the pattern was invariant (i.e., 90° rotation in the image plane), but the transformed pattern differed from the untransformed one, the subjects were required to do the same for the frame (also in the same direction).

Three dependent measures were recorded: (1) The number of arrow key presses, (2) the time elapsed between onset of the stimulus patterns and the button press, and, (3) the number of attempts needed to perform a trial correctly. As in the previous experiments, the trials that were treated incorrectly were repeated maximally twice at a random position in the rest of the block. Compression factors of 1 and .9397 were considered equal. If the frame around a pattern could not be brought into correspondence with the transformation applied to the pattern after three attempts, the trial was not repeated further and received a score 4 on the third dependent variable (the score would be 3 if there were a correct setting of the frame on the third attempt).

In this and the next experiment, the total number of trials was only half of the number in Experiments 1 and 2. However, this amounts to the same number of repeated observations per rotation-compression combination per kind of pattern (viz., three), because the previous experiments used a same-different task (which doubles the total number of trials to yield the same number of observations for the same-responses). Again, the total number of trials (i.e., 1,344) was randomly divided in 16 blocks that could be managed without a break. The first blocks by each subject took about 30 min but after some practice only about 20 min were needed. An additional subject started this experiment but did not finish it because she was excessively slower than other subjects. Her first three blocks required more than one hour each. Because the course requirement is limited to four hours, this subject was made to stop. The data were not analyzed. As in the previous experiments, the naive subjects performed one block of practice trials, similarly constituted as the experimental ones, but with different dot patterns. The rest of the method used in this experiment was analogous to that of Experiment 1, except for the fact that only two or three blocks could be taken in a 1- h session. Again, sessions were spread over several days within a one week period.

RESULTS

The data were analyzed by three ANOVAs on the three dependent measures in a within-subjects design with four orthogonal factors: Kind of patterns (2 levels: symmetric vs. random), Pattern type (7 levels: VHLR, VH, LR, V, H, L, R), Rotation (8 levels: 90°, 67.5°, 45°, 22.5°, 0°, -22.5°, -45°, and -67.5°), and Compression (4 levels: 1, .9397, .766, and .5). Because the results of these three ANOVAs were very similar, only those for the RTs associated with a correct response will be reported. Three notes are needed here. First, there were obvious exceptions to the similarity of the results for the different dependent variables, such as the fact that random patterns did not require more button presses than symmetric ones. This results from the procedure in which all frames had to be adjusted and all differed to the same degree from the patterns within the frame. Second, the RTs are the times from stimulus onset until button press. Note, however, that these include much more than the mere perceptual processing of the stimuli (e.g., the use of the arrow keys). They are, therefore, quite long and are further expressed in seconds. Third, of

TABLE 7
ANOVA on RTs for Experiment 3

Effect	*df*	*F*	*p*	%
Kind	1,3	22.57	.05	35.38
Pattern	6,18	5.52	.005	0.82
Rotation	7,21	6.48	.0005	35.76
Compression	3,9	22.09	.0005	5.24
K × P	6,18	11.79	.00001	1.81
K × R	7,21	20.32	.00001	5.23
P × R	42,126	1.77	.01	2.90

all 1,344 RT observations per subject, only a small number was not associated with a correct response but with the last incorrect one. These numbers were 13, 60, 10, and 3 for the four (pooled) subjects. Means based on excluding these "incorrect" RTs were very similar. The same principles as in Experiments 1 and 2 were followed. All significant ($p < .05$) effects are indicated in Table 7.

The main effect of *Kind*, explaining 35% of the variance, was due to the fact that much more time was needed to adjust the frame of a random dot pattern than for a symmetric one (6.2 vs. 10.1 s). The interaction with *Pattern* was caused by the large and systematic effects for the symmetric patterns, whereas the differences for the random ones were unsystematic. The differences between the types of symmetry resulted from effects of number and orientation of axes (see Table 2, third column). With respect to the former, only the difference between quadruple symmetry (5.0 s) and the others were significant. Double symmetry did not differ from symmetry about a single axis (6.4 vs. 6.3 s). With respect to the latter, VH differed from LR, as did both single main axes (V, H) from both diagonals (L, R).

The main effect of *Rotation* was of equal importance as the difference between both kinds of patterns, in terms of the explained variance. Of course, a good part of this rotation effect will be due to the procedural aspect which required more button presses in cases of larger rotation angles. However, three observations make it unlikely that there is nothing more to it. First, similar rotation effects were obtained with different procedures (cf. supra). Secondly, random and symmetric patterns produced different results although the rotation angles are the same. Third, the rotation effects were also found with the other dependent measures (e.g., errors). As in the previous cases, this rotation effect resulted from larger RTs at larger angles deviating from V, both CW and CCW (see Table 3, next two columns). Nevertheless, this rotation effect is much more pronounced for the random dot patterns and it differs somewhat for the different types of symmetric ones. For example, the quadratic component is highly reliable for all types, except for symmetry about four axes, where it is just marginally significant ($p < .06$). In addition,

TABLE 8
Effects of Compression on Response Times (in s) for Different Kinds of Patterns in Experiments 3 and 4

	Experiment 3		Experiment 4	
Compression	SYM	RND	SYM	RND
1	5.5	9.1	4.2	4.9
.9397	5.8	9.6	4.5	5.2
.766	6.4	10.2	5.4	5.8
.5	7.0	11.6	5.8	6.5

the quartic trend components are only reliable for those types of symmetry in which the V axis plays a role (VHLR, VH, and V).

The main effect of *Compression* was reliable but explained only 5% of the variance and did not interact with other factors. Longer RTs were associated with larger compressions (see Table 8, first two columns).

DISCUSSION

As in all previous experiments, random dot patterns yielded much longer RTs to perform the task than symmetric ones. Although some of the problems mentioned above still remain, some arguments can be given against the use of local cues as an explanation for the results obtained in this experiment. For example, some of the differences between the symmetries involving a V axis (VHLR, VH, and V) were reliable, in contrast with Experiment 2, where it was suggested that subjects probably only relied on local parallelism. The procedural aspect that no transformation had to be applied in cases of pattern invariance cannot account for this result, because it would hold for pure rotation of quadruple symmetry only.

Another difference between Experiment 3 and the previous ones, is the absence of proximity cues. Of course, when subjects are transforming the frame in accordance with the affine transformation causing the difference between the transformed version at the right of the screen and the original one at the left, they do not necessarily recover surface orientation. It is possible to remain within the image plane and use proximity between pattern elements and frame border to check a particular setting but it is no longer possible, as in the previous experiments, to reject some trials immediately on the basis of that local cue. In sum, more global effects due to the order of symmetry will probably have played some role in addition to the pure local cues. But it is not clear whether some real recovery of 3-D surface orientation has taken place.

Experiment 4

METHOD

Stimuli and apparatus

In each trial two polygons, each connecting 24 coordinates in accordance with

TABLE 9
ANOVA on RTs for Experiment 4

Effect	*df*	*F*	*p*	%
Kind	1,3	118.47	.005	2.63
Pattern	6,18	14.00	.00001	4.01
Rotation	7,21	35.35	.00001	40.78
Compression	3,9	12.66	.005	11.94
K × P	6,18	17.20	.00001	5.66
K × R	7,21	18.55	.00001	5.04
P × R	42,126	3.51	.00001	5.82
K × P × R	42,126	2.87	.00001	4.92
P × R × C	126,378	1.50	.005	6.98

some prespecified rules (to avoid intersections and to preserve the type of symmetry if any) were presented as in Experiment 2. In this experiment, however, the untransformed polygons were surrounded by a square frame, whereas the transformed ones were included in a square initially. As in Experiment 3 for the dot patterns, the frame could be rotated and transformed into a rectangle or a parallelogram by using the arrow keys of an extended keyboard.

Subjects, task and procedure

The rest of the method used in this experiment was completely analogous to that of Experiment 3.

RESULTS

All significant ($p < .05$) effects are indicated in Table 9. Whereas the main effect of *Kind* still reached statistical significance, it explained only a minor fraction of the variance (3%). Nevertheless, it was in the right direction: Frames around symmetric polygons required less time to be adjusted than those around random ones (5.0 vs. 5.6 s). This difference was most probably attenuated by the fact that oblique axes polygons seemed more difficult than in previous experiments. Likewise, the effect of number of axes (3.6, 5.5, and 5.1 s, for 4, 2, and 1 axes, resp.) was seriously modified by large differences between axis orientations for polygons with the same order of symmetry (see Table 2, fourth column). For example, VH differed from LR and V differed from all other single axis conditions. It is interesting to note that the same kind of differences were observed in Experiment 3, but they are much larger here.

The main effect of *Rotation* contributed most to the total systematic variance (41%). As in the previous experiments, this was largely due to increased RTs for larger CW and CCW rotation angles (see Table 3, final two columns). However, as in Experiment 3, the rotation effect differed for random polygons and different types of symmetric ones. For example, in addition to the large quadratic component, rotation yielded reliable cubic and quartic trend components for random polygons. For quadruple symmetry, only the quartic component was reliable, whereas linear and/or quartic components were added to the large quadratic effects for the other types of symmetry.

Finally, the main effect of *Compression* produced quite large effects on the RTs (12% of the variance). As in previous experiments, longer RTs were caused by larger compressions (see Table 8, final two columns). However, this main effect was modified by a reliable third-order interaction with Rotation and Pattern. In the absence of compression, increasing rotation angles cause longer RTs, except at ±45° for the symmetric polygons. Similarly, in the absence of rotation, increasing compression yielded longer RTs, especially for the symmetric polygons. In addition, the pure effects were seriously modified by changes at the different levels of the other factor. For random polygons, the quadratic rotation effect remained highly reliable at all levels of compression, but statistically significant quartic components were added to it when compression was .766 or .5. On the other hand, the rank ordering caused by compression remained almost unchanged at all rotation angles, although the differences were somewhat larger for the ±22.5° and ±45° rotations. The changes were more dramatic for the symmetric polygons. A serious flattening of the rotation effect occurred as a result of larger compression factors. Likewise, the rank ordering caused by compression differed at different rotation angles. For example, in contrast with the random polygons, ±45° rotations made the differences between the three nontrivial compressions almost disappear completely. Finally, it is clear that performance differs largely as a function of the implied surface orientation (Rotation-Compression combination) and the type of pattern on the surface.

DISCUSSION

In contrast with previous experiments, the difference between symmetric and random polygons, which is still reliable when averaged across all types of symmetry, is completely due to the symmetry types with V axes (VHLR, VH, and V). In Experiment 2, where polygons were used in a different paradigm, a similar decrease of symmetry advantage was observed for oblique axes but was much less pronounced. However, in contrast with Experiment 2, but in agreement with Experiment 3, where dot patterns were used in the same paradigm, differences between V symmetric polygons were obtained depending on the number of axes. So, the results in Experiment 4 can be seen as combining two effects which were observed separately in the two experiments, each of which contained an aspect of the combined procedure used here.

The fact that the differences between symmetries about main or oblique axes are larger with polygons than with dot patterns can be attributed to the saliency of the local parallelism between peripheral pattern elements and frame border. However, perhaps this is not the whole story. The fact that V symmetry is easier with multiple axes cannot be explained by it. If this were the only source of information subjects relied on, there would be no difference caused by number of axes.

If one would want to argue that more local parallelisms are better because it allows more checks (four in the case of VHLR, two for VH), one should, in fact, predict the reverse effect. More time would be needed to check two and four parallelisms, unless one would suppose that they are all detected in parallel across the whole pattern. However, that would amount to the very rapid detection and use of skewed symmetry in polygons, because the different local parallelisms are in skewed symmetric positions in the display. So, perhaps stronger Gestalts are helpful after all and do allow for better or easier surface-orientation recovery. The higher-order interaction specified above seems to support this suggestion.

General Discussion

The pattern of results obtained in this set of four experiments is somewhat complex and interpretation of them is hampered by potential artifactual cues. The general strategy followed in discussing the experimental findings has been to attribute specific results to artifacts if possible, even if they would be congruent with the basic hypothesis being tested, namely that SS might be used to recover surface orientation. However, some results appear to suggest that global aspects such as order of symmetry do play a role in enhancing the ease with which orientation in depth can be recovered from SS.

In Experiments 1 and 2 it was perfectly possible to judge whether the suggested 3-D orientations of pattern and frame around it were the same or different by using local cues in the image plane. Identity could be determined in some types of symmetry on the basis of local parallelism between virtual or real lines connecting peripheral symmetrically positioned pattern elements and frame border. On the other hand, difference could be decided very quickly in some cases of random patterns by detecting close proximity between peripheral elements and frame border.

Several specific results supported this interpretation. For example, for symmetric patterns *Same* responses were generally faster than *Different* ones, whereas the reverse occurred for random patterns. Evidence for the use of local parallelism was found in the difference caused by orientation of the axes of symmetry and by the more pronounced effects for polygons versus dot patterns. The use of proximity for random patterns was restricted to particular rotation-compression combinations.

In Experiments 3 and 4, subjects were required to set the parameters of the affine transformation of the frame such that the suggested 3-D orientation would be congruent with the one implied by the affine transformed dot patterns or polygons. Using this procedure, some results were obtained that could not be attributed to local cues in the image plane only. The main finding in this respect was the advantage of multiple axes of symmetry. This is not consonant with the use of local parallelism only, unless it would be checked simultaneously at different spatial locations in the image, which would, in fact, require very fast detection of SS and surface-orientation recovery to determine that the SS positions correspond with BS ones in 3-D space. Additional support for a kind of depth recovery was obtained in a higher-order interaction indicating that specific rotation-compression combinations were easier than others.

In sum, an interpretation which could capture all of these findings seems to be that, in general, subjects cannot recover the 3-D orientation from affine transformed dot patterns or polygons very easily. As a result of this, subjects learn to detect some local cues in the image plane which can be used as a strategy to perform the task. Only a restricted number of highly regular patterns (VHLR and VH, foremost polygons) are immediately and effortlessly perceived as being oriented in depth. In these cases where the task can be performed on the basis of this preattentive detection of SS-as-BS-in-depth (to paraphrase Epstein's conceptualization of similar problems), one does not have to rely on local parallelism or proximity in the image plane.

Modes of Using NAPs

DETECTION-USE/IMPLICIT-EXPLICIT/ACTIVE-PASSIVE/ROUTES TO OBJECT RECOGNITION

In the introduction to the experiments above, I distinguished between evidence for the detection of NAPs (which is available to a certain extent) and evidence for their use (which was hitherto almost completely lacking, except for some casual observations). In addition to this logical distinction between ways of processing visual information (the latter requires the former), I noticed that particular experimental tasks can put different requirements to the visual system. For example, the experiments described under the heading of "SS as a NAP", investigated whether symmetry (SS in the more general case of orthographic projection) is helpful to determine the equivalence of patterns under affine transformations. The underlying assumption was that *Same/Different* tasks are known to be facilitated by Gestalt factors such as BS and that a similar positive effect for the case of SS might be interpreted as evidence for the perception of SS as BS-in-depth. In the experiments described under the heading of "Shape from SS", more demanding tasks were designed in that subjects had to recover the 3-D orientation from the SS in the 2-D image in order to be able to (passively) tell whether that of the frame corresponded with that of the patterns enclosed by it or to bring both (actively) in agreement.

In fact, this distinction corresponds to Lowe's (1985) suggestion of two possible routes to object recognition, one via the intermediate recovery of viewer-dependent aspects of shape, such as distance and orientation, represented in a pointwise depth map such as Marr's (1982) 2 ½-D sketch, and a second one, directly from 2-D image properties to 3-D scene characteristics, without intermediate representations. The latter route corresponds to the use of NAPs, whereas the former is dependent on different "shape-from" recovery algorithms. In the experiments summarized in Part II, the less demanding, more implicit use of SS was tested, whereas Experiments 1-4, described in full in Part III, tested the usefulness of SS in a more demanding, explicit way. Although this distinction between two routes was not very sharp in Lowe's original thesis (see his Figure 1.2 on p. 12), it appeared more vividly so in his later papers (e.g., 1987a; compare with his Figure 1 on p. 357), probably to stress his own contribution in contrast with Marr's (1982). A general notion to capture these subtle differences between ways of processing or using particular sources of visual information is *modes*, as introduced by Pick and Saltzman (1978). In their terms, "a different mode of perceiving may be implied when one type of information rather than another is extracted from a given pattern of stimulation" (Pick & Saltzman, 1978, p. 2). In this sense, the processing of SS provides a clear example of different modes being involved: SS allows an interpretation as BS-in-depth (SS as NAP) as well as the recovery of (constraints on) accurate 3-D surface orientation (shape from SS).

QUALITATIVE AFFINE INVARIANTS

The distinctions between detection and use, as well as between the different uses, were known from the outset. Yet, a third source of information being extracted from the stimulus displays, a third mode of processing these dot patterns and polygons, so to speak, was discovered in the experiments presented above. Although large and systematic advantages of symmetry were observed in all experiments, SS was only interpreted as BS-in-depth and used to recover 3-D surface orientation in the cases of closed polygons and highly regular dot patterns. In all other cases, subjects seemed to rely on local groupings or regularities such as parallelisms or proximities. These provided an additional source of information in the case of symmetric patterns (e.g., alignment between pattern border and frame). Even in the case of random patterns, this kind of local grouping occurred to provide what could be called "alignment keys" (Ullman, 1989), anchor points that can be used to find a correspondence between affine equivalent patterns. For example, some dot patterns contained quite salient curvilinear structures afforded by relatively close spacing and smooth transitions or rather special local areas caused by accidental differences in densities (e.g., clusters). Likewise, some polygons contained obvious peculiarities such as deep concavities, zigzag structures, or large, uninterrupted parts. In symmetric patterns, these local structures, randomly generated by the dot localization process, get multiplied with increased saliency as a result. In a sense, these structures or groups, generated by early perceptual grouping processes and afforded by accidental feature localizations, can be regarded as NAPs: If they are present in a pattern, they are likely to occur in affine transformed versions of the pattern resulting from perceiving the pattern from a different angle. Although not explicitly manipulated and tested in these experiments, I believe this third mode of using NAPs to be a psychologically very plausible one.

The big difference between the third mode of using NAPs and the other two is that the latter require some kind of (implicit or explicit) 3-D recovery, whereas the former allows processing to remain in the 2-D image plane. A recent exchange in the literature about *structure-from-motion*, also known as the kinetic depth effect or KDE, clarifies this (Braunstein & Todd, 1990; Sperling, Dosher, & Landy, 1990; Sperling, Landy, Dosher, & Perkins, 1989). These authors distinguish between KDE, KDE-alternative, and artifactual computations. In contrast with real KDE, which requires the computation of 3-D depth values, the KDE-alternative computation derives the response directly from the 2-D optic flow, without an intermediate stage of perceived 3-D depth, whereas an artifactual computation uses incidental stimulus or motion cues from only a small part of the stimulus. In light of the richness of normal visual information and the relativity of the artifact-information distinction (it depends on what one is interested in), it remains unclear whether these qualitative perceptual structures deserve to be called artifactual or not. I think it is useful visual information because it allows subjects to perform the task without having to recover depth, which proves slow and cumbersome in dot patterns without high-order symmetries.

Doubts about the more demanding use of NAPs or visual sources of information in recovery situations have been formulated previously. For example, with respect to computer vision, Lowe (1987a, p. 358) stated that "one difficulty with the methods for depth reconstruction is that the required inputs are often unavailable or require an unacceptably long interval of time to obtain". In a similar vain, Barnard (1984, p. 17) concluded that "there is, in general, insufficient information in a single image to construct iconic, viewer-centered representations of physical surface properties". With respect to human perception, Todd and Akerstrom (1987, p. 254) stressed that "it is important to keep in mind that ... the most commonly used properties for describing surfaces - namely, depth, orientation, and curvature - were all originally invented by geometers and need not be accepted by default as the appropriate descriptors for perceptual psychology". They refer to Gibson (1979) for his alternative ways of describing surface layout that do not require a point-by-point estimate of depth and/or orientation (e.g., concavities and convexities).

However, the role of image regularities such as parallelisms, proximities, curvilinearities, clusters, and concavities, has been noticed only recently. I propose to call these structures *qualitative affine invariants*, to indicate that they are preserved under affine transformation only in a qualitative sense. For example, exact curvature is not affine invariant, but the sign of curvature (concave or convex) is. Likewise, interelement distances are not exactly identical when viewed from different angles, but relative densities remain quite comparable. The experiments introduced above suggest that human intermediate level vision is characterized by its qualitative and opportunistic nature, ignoring precise geometric measurements and making use of whatever "features" appear in an image. Moreover, these "features" need not be absolute invariants under all projections of the object giving rise to them. By relying on different regular groups simultaneously, the degree of certainty of the inference from each individual one does not have to be perfect (for a computational approach, see Van Gool, Wagemans, Vandeneede, & Oosterlinck, 1990).

Only very recently (although foreshadowed by an important paper by Witkin & Tenenbaum, 1983), a similar trend towards qualitative vision can be noticed in computational approaches to regularity detection and surface-orientation recovery. For example, at the "AAAI Workshop on Qualitative Vision" held in Boston at the end of July 1990, computational work was presented about qualitative features and their relationships (Jacobs, 1990) and evaluation of perceptual groups (Kahn, Winkler, & Chong, 1990). With respect to recovery, it was shown that partial qualitative shape information can be obtained from stereo without going through a detailed disparity map (Edelman, 1990). Likewise, Weinshall (1990) demonstrated that the sign of the Gaussian curvature (but not the Gaussian curvature itself) and the qualitative direction of motion (but not the exact value) can be computed directly from the local motion disparity field. This corresponds nicely with recent theoretical and psychophysical work about more qualitative, nonmetric structural shape descriptors, such as nominal (e.g., Koenderink, 1984; Koenderink & van Doorn, 1982) or ordinal structure (e.g., Todd & Reichel, 1989, 1990).

Conclusion and Future Research Directions

Based on current approaches and available findings, I think it makes good sense to distinguish between different modes of detecting and using image regularities. In addition to the commonly made distinction between two routes to object recognition (via NAPs or via an intermediate depth map), a third mode, which uses the qualitative affine invariants as such, seems to occur in human and machine vision. Although these are regarded as artifactual when one is interested in depth recovery only, they might well be the most often used available information in situations where 3-D surface orientation is not salient, such as in relatively sparse dot patterns.

Two lines of future research seem to be most fruitful in this respect. First of all, it would be interesting to test the usefulness of NAPs such as SS with less fragmented and arbitrary stimuli in less demanding situations. For example, apparent motion between affine equivalent closed contours (with or without SS) might be a viable paradigm. Secondly, qualitative use of 2-D image regularities such as parallelism, curvilinearities, and concavities deserves to be studied in more detail. In light of some recent computational approaches to extract these properties (e.g., Ahuja & Tuceryan, 1989; Lowe, 1989; Sha'ashua & Ullman, 1988; Smits & Vos, 1986), it might well be possible to manipulate their presence more explicitly. Moreover, in light of the opportunistic and flexible nature of human vision, it would be especially interesting to investigate how these qualitative affine invariants interact (e.g., a curvilinearity becomes more salient when repeated in symmetric or parallel positions). Although classic Gestalt approaches have tackled some of these phenomena long ago, I think the relevance of the Gestalt laws of perceptual organization to so-called higher visual tasks such as object recognition and scene perception remains largely to be discovered and specified.

The author is supported by a grant from the Belgian National Fund for Scientific Research (NFWO). The research benefits from interdisciplinary collaboration with the Department of Electronic Engineering (ESAT-MI2) and the Department of Mathematics, which is sponsored by the University of Leuven (OT/89/11), the Belgian Fund for Joint Basic Research (FKFO), and ESPRIT-BRA-6448 (VIVA). I gratefully acknowledge the discussions with G. d'Ydewalle, P. Vanden Bossche, and L. Van Gool through all stages of the research; the technical support by D. Christiaens, A. De Troy, M. Lenaerts, V. Swinnen, P. Vanden Bossche, and J. Van Rensbergen; the assistance in running part of the subjects and part of the data analyses by E. Ver Eecke and A. Victoir; and the comments on a previous version of the manuscript by G. d'Ydewalle, K. Humphrey, P. Jolicoeur, and J. Todd.

Correspondence should be addressed to Johan Wagemans, Laboratory of Experimental Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium (e-mail: fpaas10@blekul11.bitnet).

References

Ahuja, N., & Tuceryan, M. (1989). Extraction of early perceptual structure in dot patterns: Integrating region, boundary, and component Gestalt. *Computer Vision, Graphics, and Image Processing, 48*, 304-356.

Attneave, F. (1972). Representation of physical space. In A. W. Melton, & E. Martin (Eds.), *Coding processes in human memory* (pp. 283-306). New York: Wiley.

Attneave, F. (1982). Prægnanz and soap bubble systems: A theoretical exploration. In J. Beck (Ed.), *Organization and representation in perception* (pp. 11-29). Hillsdale, NJ: Erlbaum.

Attneave, F., & Frost, R. (1969). The determination of perceived tridimensional orientation by minimum criteria. *Perception & Psychophysics, 6*, 391-396.

Bagnara, S., Boles, D.B., Simion, F., & Umilta, C. (1983). Symmetry and similarity effects in the comparison of visual patterns. *Perception & Psychophysics, 34*, 578-584.

Barnard, S.T. (1984). *An inductive approach to figural perception*. (Tech. Note No. 325). Menlo Park, CA: SRI International, AI Center.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115-147.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20*, 38-64.

Biederman, I., & Shiffrar, M.M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental psychology: Learning, Memory, and Cognition, 13*, 640-645.

Braunstein, M.L., & Todd, J.T. (1990). On the distinction between artifacts and information. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 211-216.

Brunswik, E. (1952). *Systematic and representative design of psychological experiments*. Berkeley, CA: University of California Press.

Cutting, J.E. (1987). Rigidity in cinema seen from the front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 323-334.

Cutting, J.E. (1988). Affine distortions of pictorial space: Some predictions for Goldstein (1987) that La Gournerie (1859) might have made. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 305-311.

Cutting, J.E., & Millard, R.T. (1984). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology: General, 113*, 198-216.

Edelman, S. (1990). Local qualitative shape from stereo without detailed correspondence. *Proceedings of the AAAI Workshop on Qualitative Vision* (pp. 101-105). Boston, MA: AAAI Society.

Enns, J.T. (1990). Three-dimensional features that pop-out in visual search. In D. Brogan (Ed.), *Visual search* (pp. 37-45). Hants, UK: Taylor & Francis.

Enns, J.T. (1992). Sensitivity of early human vision to 3-D orientation in line-drawings. *Canadian Journal of Psychology*, this issue.

Enns, J.T., & Rensink, R.A. (1990a). Influence of scene-based properties on visual search. *Science, 247*, 721-723.

Enns, J.T., & Rensink, R.A. (1990b). Sensitivity to three-dimensional orientation in visual search. *Psychological Science, 1*, 323-326.

Epstein, W. (1973). The process of "taking-into-account" in visual perception. *Perception, 2*, 267-285.

Epstein, W., & Babler, T. (1989). Perception of slant-in-depth is automatic. *Perception & Psychophysics, 45*, 31-33.

Epstein, W., & Babler, T. (1990). In search of depth. *Perception & Psychophysics, 48*, 68-76.

Epstein, W., & Broota, K.D. (1986). Automatic and attentional components in perception of size-at-a-distance. *Perception & Psychophysics, 40*, 256-262.

Epstein, W., Hatfield, G., & Muise, G. (1977). Perceived shape at a slant as a function of processing time and processing load. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 473-483.

Epstein, W., & Lovitts, B.E. (1985). Automatic and attentional components in perception of shape-at-a-slant. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 355-366.

Foster, D.H., & Ferraro, M. (1989). Visual gap and offset discrimination and its relation to categorical identification in brief line-element displays. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 771-784.

Fox, J. (1975). The use of structural diagnostics in recognition. *Journal of Experimental Psychology: Human Perception and Performance, 104*, 57-67.

Friedberg, S.A. (1986). Finding the axes of skewed symmetry. *Computer Vision, Graphics, and Image Processing, 34*, 138-155.

Garner, W.R. (1974). *The processing of information and structure.* Potomac, ML: Erlbaum.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin.

Hakalahti, H. (1983). *Estimation of surface orientation using the hypothesis of object symmetry* (TR-1324). Maryland, MD: Center for Automation Research.

Ittelson, W.H. (1952). *The Ames demonstrations in perception.* Princeton, NJ: Princeton University Press.

Ittelson, W.H. (1960). *Visual space perception.* New York: Springer.

Jacobs, D.W. (1990). Opportunistic recognition. *Proceedings of the AAAI Workshop on Qualitative Vision* (pp. 147-151). Boston, MA: AAAI Society.

Kahn, P., Winkler, A., & Chong, C.Y. (1990). Qualitative evaluation of perceptual groups. *Proceedings of the AAAI Workshop on Qualitative Vision* (pp. 157-161). Boston, MA: AAAI Society.

Kanade, T. (1981). Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence, 17*, 409-460.

Kanade, T., & Kender, J.R. (1983). Mapping image properties into shape constraints: Skewed symmetry, affine-transformable patterns, and the shape-from-texture paradigm. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (Vol. 1, pp. 237-257). New York: Academic Press.

King, M., Meyer, G.E., Tangney, J., & Biederman, I. (1976). Shape constancy and a perceptual bias towards symmetry. *Perception & Psychophysics, 19*, 129-136.

Koenderink, J. J. (1984). What does the occluding contour tell us about solid shape? *Perception, 13*, 321-330.

Koenderink, J. J., & van Doorn, A. J. (1982). The shape of smooth objects and the way contours end. *Perception, 11*, 129-137.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt-Brace.

Lowe, D.G. (1985). *Perceptual organization and visual recognition* (PhD Thesis, Department of Computer Science, Stanford University). Boston, MA: Kluwer.

Lowe, D.G. (1987a). Three-dimensional object recognition from two-dimensional images. *Artificial Intelligence, 31*, 355-395.

Lowe, D.G. (1987b). The viewpoint consistency constraint. *International Journal of Computer Vision, 1*, 57-72.

Lowe, D.G. (1989). Organization of smooth image curves at multiple scales. *International Journal of Computer Vision, 3*, 119-130.

Lowe, D.G. (1990). Visual recognition as probabilistic inference from spatial relations. In A. Blake & T. Troscianko (Eds.), *AI and the eye* (pp. 261-279). Chichester, UK: Wiley.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* San Francisco, CA: Freeman.

Palmer, S.E. (1983). The psychology of perceptual organization: A transformational approach. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (Vol. 1, pp. 269-339). New York: Academic Press.

Pick, H.L., Jr., & Saltzman, E. (1978). Modes of perceiving and processing information. In H. L. Pick, Jr., & E. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 1-20). Hillsdale, NJ: Erlbaum.

Pomerantz, J.R. (1977). Pattern goodness and speed of encoding. *Memory & Cognition, 5*, 235-241.

Rock, I. (1977). In defense of unconscious inference. In W. Epstein (Ed.), *Stability and constancy in visual perception: Mechanisms and processes* (pp. 321-373). New York, NY: Wiley.

Rock, I. (1983). *The logic of perception.* Cambridge, MA: MIT Press.

Rock, I., & DiVita, J. (1987). A case of viewer-centered object perception. *Cognitive Psychology, 19*, 280-293.

Sha'ashua, A., & Ullman, S. (1988). Structural saliency: The detection of globally salient structures using a locally connected network. *Proceedings of the Second International Conference on Computer Vision* (pp. 321-327). Tampa, FL: IEEE Computer Society.

Shepard, R.N. (1981). Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 279-341). Hillsdale, NJ: Erlbaum.

Smits, J.T.S., & Vos, P.G. (1986). A model for the perception of curves in dot figures: The role of local salience of "virtual lines". *Biological Cybernetics, 54*, 407-416.

Sperling, G., Dosher, B.A., & Landy, M.S. (1990). How to study the kinetic depth effect experimentally. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 445-450.

Sperling, G., Landy, M.S., Dosher, B.A., & Perkins, M.E. (1989). Kinetic depth effect and identification of shape. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 826-840.

Stevens, K.A. (1980). *Surface perception from local analysis of texture and contour* (PhD. Thesis, Tech. Rep. 512). Cambridge, MA: MIT AI Lab.

Stevens, K.A. (1983). Surface tilt (the direction of slant): A neglected psychophysical variable. *Perception & Psychophysics, 33*, 241-250.

Todd, J. T., & Akerstrom, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 242-255.

Todd, J.T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics, 48*, 419-430.

Todd, J.T., & Reichel, F.D. (1989). Ordinal structure in the visual perception and cognition of smoothly curved surfaces. *Psychological Review, 96*, 643-657.

Todd, J.T., & Reichel, F.D. (1990). Visual perception of smoothly curved surfaces from doubly-projected contour patterns. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 665-674.

Treisman, A. (1985). Preattentive processes in vision. *Computer Vision, Graphics, and Image Processing, 31*, 156-177.

Treisman, A. (1988). Features and objects: The 14th 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.

Treisman, A., & Patterson, R. (1984). Emergent features, attention, and object perception. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 12-31.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 459-478.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition, 32*, 193-254.

Van Gool, L. Wagemans, J., Vandeneede, J., & Oosterlinck, A. (1990). Similarity extraction and modelling. *Proceedings of the Third International Conference on Computer Vision* (pp. 530-534). Osaka, Japan: IEEE Computer Society.

Wagemans, J. (1988). Modules in vision: A case study of interdisciplinarity in cognitive science. *Acta Psychologica, 67*, 59-93.

Wagemans, J. (1990). "Smart" mechanisms emerging from cooperation and competition between modules. *Psychological Research, 52*, 181-196.

Wagemans, J. (1992). Skewed symmetry: A nonaccidental property used to perceive visual form. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2).

Wagemans, J., Van Gool, L., & d'Ydewalle, G. (1990). Visual search in dot-patterns with bilateral and skewed symmetry. In D. Brogan (Ed.), *Visual search* (pp. 99-114). Hants, UK: Taylor and Francis.

Wagemans, J., Van Gool, L., & d'Ydewalle, G. (1991). Detection of symmetry in tachistoscopically presented dot patterns: Effects of multiple axes and skewing. *Perception & Psychophysics, 50*, 413-427.

Wagemans, J., Van Gool, L., & d'Ydewalle, G. (1992). Orientational effects and component processes in symmetry detection. *Quarterly Journal of Experimental Psychology, 44A*, 475-508.

Weinshall, D. (1990). Shortcuts in the computation of qualitative 3-D shape and motion invariants. *Proceedings of the AAAI Workshop on Qualitative Vision* (pp. 31-35). Boston, MA: AAAI Society.

Witkin, A.P., & Tenenbaum, J.M. (1983). On the role of structure in vision. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (Vol. 1, pp. 481-543). New York: Academic Press.