PERGAMON

Vision Research 39 (1999) 4244–4257

Vision Research

www.elsevier.com/locate/visres

The role of convexity in perceptual completion: beyond good continuation

Zili Liu [a,*], David W. Jacobs [b], Ronen Basri [c]

[a] *Department of Psychology, Rutgers University, Newark, NJ 07102, USA*
[b] *NEC Research Institute, 4 Independence Way, Princeton, NJ 08540, USA*
[c] *Department of Applied Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel*

Received 7 October 1997; received in revised form 7 June 1999

Abstract

Since the seminal work of the Gestalt psychologists, there has been great interest in understanding what factors determine the perceptual organization of images. While the Gestaltists demonstrated the significance of grouping cues such as similarity, proximity and good continuation, it has not been well understood whether their catalog of grouping cues is complete — in part due to the paucity of effective methodologies for examining the significance of various grouping cues. We describe a novel, objective method to study perceptual grouping of planar regions separated by an occluder. We demonstrate that the stronger the grouping between two such regions, the harder it will be to resolve their relative stereoscopic depth. We use this new method to call into question many existing theories of perceptual completion (Ullman, S. (1976). *Biological Cybernetics*, *25*, 1–6; Shashua, A., & Ullman, S. (1988). 2nd International Conference on Computer Vision (pp. 321–327); Parent, P., & Zucker, S. (1989). *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*, 823–839; Kellman, P. J., & Shipley, T. F. (1991). *Cognitive psychology*, Liveright, New York; Heitger, R., & von der Heydt, R. (1993). A computational model of neural contour processing, figure-ground segregation and illusory contours. In *Internal Conference Computer Vision* (pp. 32–40); Mumford, D. (1994). *Algebraic geometry and its applications*, Springer, New York; Williams, L. R., & Jacobs, D .W. (1997). *Neural Computation*, *9*, 837–858) that are based on Gestalt grouping cues by demonstrating that convexity plays a strong role in perceptual completion. In some cases convexity dominates the effects of the well known Gestalt cue of good continuation. While convexity has been known to play a role in figure/ground segmentation (Rubin, 1927; Kanizsa & Gerbino, 1976), this is the first demonstration of its importance in perceptual completion.

Keywords: Amodal completion; Grouping; Convexity; Good continuation; Stereoscopic depth

1. Introduction

In studying perceptual organization, we consider under what circumstances a set of image elements will be perceived as a single unit. One of the most basic questions we can ask, then, is given two image elements, how well or poorly do they group into a single entity? In this paper, we describe work in which two elements are separated by a possible occluder, and consider the extent to which these elements may be perceived as joining into a single shape underneath the occluder. This is often termed *amodal completion*)[1].

When two fragments group together into an object, the visual system perceives part of the boundary of each fragment as due to the true boundary of the object, and part of the boundary as artificial, belonging instead, e.g. to the boundary of an intervening, occluding object. The visual system hypothesizes that at some point,

* Corresponding author. Tel.: + 1-973-353-1917; fax: + 1-973-353-1171.

E-mail address: liu@psychology.rutgers.edu (Z. Liu)

[1] In this paper, we use the terms 'perceptual completion' and 'perceptual grouping' interchangeably. We note, however, that perceptual completion entails more than grouping. For instance, when dots move with a common velocity, they are grouped together by the Gestalt law of 'common fate'. However, there is no perceptual completion here as the dots are not perceptually connected into a surface. On the other hand, in amodal completion, the separated elements are perceptually connected by a surface behind an occluder.

0042-6989/99/$ - see front matter
PII: S0042-6989(99)00141-8

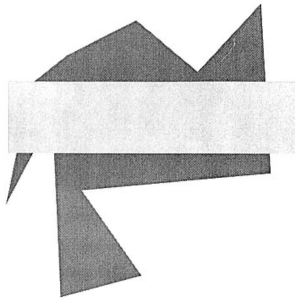


Fig. 1. Elements on either side of an occluder appear as if they may complete into a single entity in the image.

the boundary continues in a way that may not be apparent from image brightness contrast, connecting each fragment to the other, as shown in Fig. 1. In amodal completion, the location at which the suspected object boundary first becomes invisible is typically indicated by occlusion cues. A comparable process occurs in *modal completion*, when specific boundaries are perceived as illusory contours. We call these completing portions of boundary that are not apparent in the image from intensity changes *connecting curves*.

It has been hypothesized that the relative position of the beginning and end of a *connecting curve*, and the shape of the visible boundary at these points, plays a key role in our percept of amodal completion (e.g. Kanizsa, 1979). Much effort has focused on applying the Gestalt grouping cue of good continuation to explain both amodal and modal completion. These efforts argue that the smoothness of the possible *connecting curves* between object fragments determines how well these fragments group together. Computational and psychological models have provided explicit theories of how to measure this smoothness.

In this paper, we consider the role that convexity, another Gestalt cue, plays in amodal completion. While it has been shown that convexity can be a powerful cue in figure/ground determination (Rubin, 1927; Kanizsa & Gerbino, 1976; Gibson, 1994; Driver & Baylis, 1995; Bertamini & Friedenberg, 1999), convexity has not been shown to play a role in grouping, and is not present in most current models of perceptual grouping. We propose a taxonomy of convexity relationships, based on the work of Jacobs (1996), and suggest how these relations might play a role in grouping. We contrast the possible role of convexity relations with two important models of good continuation based on *relatability* and *curves of least energy*. We show that models that assign a role to convexity can make quite different predictions about the strength of amodal completions than do these good continuation models.

To test these models, it is necessary to have a mechanism for determining the strength of an amodal completion. Consequently, we have devised a novel method of doing this based on performance of an objective stereoacuity task. We show that amodal completion impedes a subject's ability to determine whether two planar fragments are coplanar. That is, perceptual grouping creates a bias to see the grouped object fragments as coplanar. This is in line with prior results (Mitchison & Westheimer, 1984), but our explanation of the phenomenon and our application of it to judging the strength of amodal completions is novel.

With this test in hand, we are able to evaluate the role of convexity in grouping by examining the strength of amodal completion in situations in which models based solely on good continuation will make different predictions from those that assign a role to convexity. We show that for stimuli in which good continuation models make no prediction, convexity can 'break the tie'. More significantly, we show that convexity cues can even dominate good continuation cues. These experiments demonstrate that theories of perceptual grouping must take into account the role of convexity.

1.1. Models of completion

Clearly, many different types of grouping cues can play a role in amodal completion, including symmetry, texture, and proximity. However, in this paper we consider cases in which those cues are held constant, focusing on the role of the shape of the *connecting curves*, and their relation to the object fragments they join. Mainly, theories that address this issue focus on the way that *connecting curves* form good continuations. We begin with a brief review of models of good continuation. Then we review some of our own past work, which focused instead on the role of convexity. Finally, we show how these theories can be used to make specific predictions of human amodal completion.

Psychophysical demonstrations that human perception prefers organizations of contours that are smooth dates back to the Gestalt psychologists, and is discussed, for example, in Köhler (1929), Kanizsa (1979) and Rock and Palmer (1990). Precise models of how to quantify this smoothness, or good continuation, are more recent. We first focus on models that consider the energy of the curve, then discuss the *relatability* idea of Kellman and Shipley (1991).

Ullman (1976) first proposed a specific model of perceptual smoothness. Ullman notes that where a possible *connecting curve* between two elements begins and ends, the visible boundary of each element specifies the tangent of the *connecting curve*, if it is to smoothly connect to the visible boundary. This leads one to conceptualize the problem of determining the shape of an illusory contour as a problem of interpolating a smooth *connecting curve*, given elements that specify the position and tangent directions of the two endpoints of the curve. Ullman suggested that illusory contours are

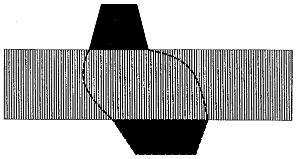


Fig. 2. Two elements are separated by a possible occluder. The dashed curves indicate two possible *connecting curves* joining them.

filled-in by a pair of circular arcs that smoothly join and minimize the integral of squared curvature.

Inspired by Ullman (1976) and Brady and Grimson (1981), Horn (1981) then proposed the mathematically more elegant idea that perceptual completion is based on finding the *connecting curve* of least energy. We may define the energy of a curve as follows. Let Γ be a curve parameterized by s, so that $\Gamma(s)$ denotes a point on the curve. Let $\kappa(s)$ denote the curvature of Γ at that point. Then define the energy, E, of the curve by:

$$E(\Gamma) = \int_{\Gamma(s_0)}^{\Gamma(s_1)} (\kappa^2(s) + \lambda)\mathrm{d}s \qquad (1)$$

where s_0 is the starting point of the curve, and s_1 is its ending point. Here λ is a parameter that makes a contribution to the energy linearly proportional to the length of the contour. Horn (1981) defines the curve of least energy as the curve that connects two points, has the appropriate tangent direction at those points, and minimizes Eq. (1)[2]. Horn (1981) then discusses some properties of the curve of least energy. Mumford (1994) discusses the history of these curves, attributing their first analysis to Euler.

Mumford further develops properties of these minimum energy curves, relating them to a generative model of curve formation. Curves of least energy and their relatives are used in a number of computational models of perceptual grouping, such as Shashua and Ullman (1988), Parent and Zucker (1989), Nitzburg and Mumford (1990), Trytten and Tuceryan (1991), Williams and Hanson (1996) and Williams and Jacobs (1997). Some of this work has spurred further psychophysics (e.g. Field, Hayes & Hess, 1993) that provides further evidence for the role of smoothness in perceptual grouping.

The above work primarily focused on perceptual organization effects such as illusory contour formation and judgments of curve salience. We now wish to use this work to make specific predictions about human performance in amodal completion. Specifically, we consider the question of how strongly two image elements, separated by a possible occluder, group together. The first model we consider will state that, when other factors are held constant, the strength of grouping will be inversely proportional to the energy of the curves of least energy that can serve as two *connecting curves* between the elements (see Fig. 2). We call this hypothesis the *energy model*, and consider it to be a straightforward application of existing ideas to the problem of amodal completion.

This model can be used to make predictions about the relative strength of two possible completions. Suppose that two pairs of image elements are each separated by an occluder. The energy model predicts that human subjects will prefer to form a single organization using the pair of elements that has a lower energy pair of *connecting curves*. This theory is not yet precise, since the energy of a curve depends on a free parameter λ. λ controls the trade-off between the length of the *connecting curve* and its total amount of curvature. So in our experiments rather than using the energy directly, we rely on only two general assumptions. First, we assume that if the distance between two endpoints is held fixed, but their tangent directions are rotated so that they point further away from each other, a connection between them will be weaker. That is, the greater the curvature needed to connect the curves, the weaker is a completion involving them, all other factors being equal. Second, that if the distance between two end points increases while their configuration is otherwise held constant, there will also be weaker completion. These assumptions seem to directly embody the factors motivating the energy model, and will be true for almost all conditions and possible versions of the energy model.

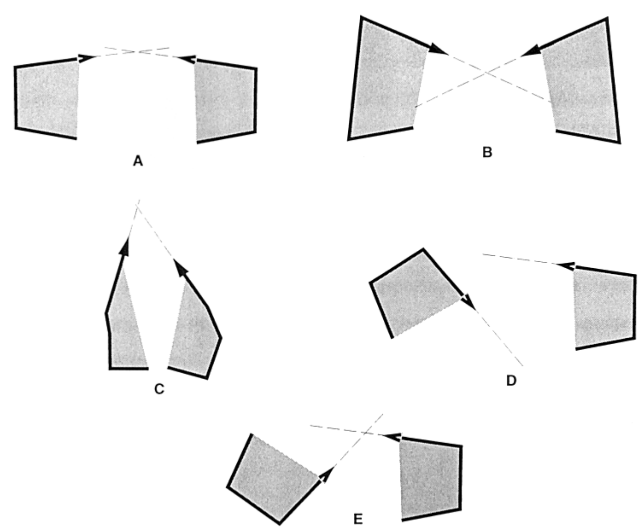


Fig. 3. Examples of relatability and non-relatability. Image elements are shown shaded, with their visible boundary shown as dark lines. Arrows indicate the position and tangent direction of an endpoint of this visible boundary. Dashed lines show possible extensions of these. In (A) and (B) the dashed lines are relatable. In (C), they are not relatable because they form an acute angle (more than 90° must be turned in joining them). In (D), they are not relatable because they do not intersect. In (E), they are not relatable because they join reversing figure and background.

[2] Horn considers this equation without λ.

Kellman and Shipley (1991) have articulated a second specific model of good continuation, called *relatability*. To define relatability, recall that we assume that the visible boundary specifies the position and tangent direction of the endpoints of any possible *connecting curve*. Imagine extending each of these visible boundaries along a straight line, in the direction of the tangent at the endpoint. Relatability makes use of one additional fact about these endpoints: the side of this extending line that contains the figure, and the side which is background. Then two endpoints will be relatable if these extensions intersect, if their angle of intersection is less than 90°, and if they intersect with no reversal as to which side the figure lies on. See Fig. 3.

A relatability model of amodal completion will predict that, if other factors are held equal, elements with relatable connections will form stronger amodal completions than elements without. Kellman and Shipley (1991) present considerable evidence for this model. We note that relatability is somewhat akin to the energy model. It is straightforward to show that two edges are relatable if and only if there exists a *connecting curve* that preserves a constant sense of figure/ground, has either always positive or always negative curvature, and such that the integral of the absolute value of its curvature is less than 90°. This last condition means that in many cases relatable edges also have *connecting curves* with lower energy than non-relatable edges, although this need not always be the case. Relatability differs clearly from energy models in explicitly penalizing inflections[3]. If all *connecting curves* between two edges contain an inflection, they are not relatable, although inflections need not greatly increase the energy of a curve.

Finally, we derive a third model based on the convexity relationships between two figural elements. This model is specifically based on a computational grouping method described in Jacobs (1996). However, we should stress that others, such as Kanizsa and Gerbino (1976) have emphasized the role of convexity in figure/ground discrimination, and Hoffman and Richards (1984) have presented compelling theoretical and psychophysical evidence for the role of convexity in determining the parts of objects. Our *convexity* model predicts that the type of convexity relationship between two elements will play a role in determining the strength of amodal completion. So we begin by presenting a taxonomy of these possible relationships.

Recall that we assume that each element has some fraction of its boundary visible, and some fraction being occluded. We define the *convex extension* of an element to be the maximal region that includes the element without introducing any additional concavities in its boundary. Intuitively, the convex extension of an element is the region enclosed by choosing *connecting curves* that are straight lines specified by the position and tangent direction of the endpoints. If these lines intersect, the enclosed region is finite, otherwise it is infinite. See Fig. 4 for an illustration. Jacobs (1996) divides the possible relationships between two elements into three classes: type I, type II and type III. A type I relation is present if the convex extension of each element includes the endpoints of the other's visible boundary. In this case, it is possible to join the elements with convex *connecting curves*. If the elements are each convex to begin with, it will be possible to join them into a single convex shape if and only if they have a type I relationship. A type II relation occurs when the convex extensions of the elements intersect, but do not include the endpoints of each contour. In this case, the boundary endpoints may belong to adjacent convex parts of an object, but not to the same convex part. If the convex extensions do not intersect, we call this a type III relationship. In this case the contour endpoints may not even belong to adjacent convex parts of the object. In addition to these types, we introduce a new one, called type I.5. This occurs when the convex extension of one contour includes the endpoints of the other, but not vice versa. This occurs when the two contours may be adjacent convex parts, and when one may be a bump, or protrusion on the other. The convexity model will predict that convexity type will play a role in the strength of amodal completion, with lower types associated with preferred organizations.

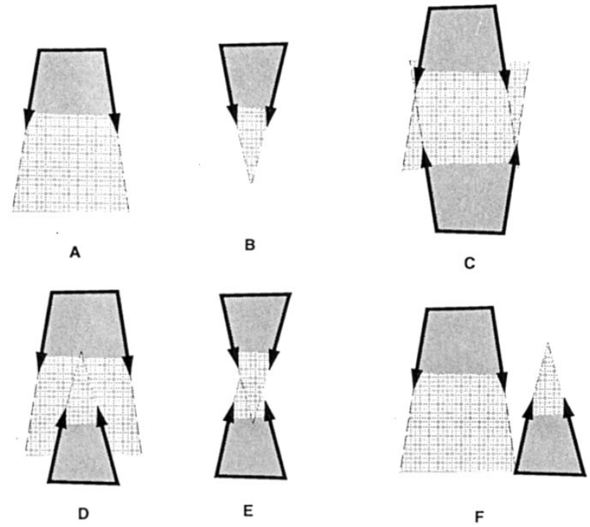


Fig. 4. In (A) and (B), the cross-hatched region illustrates the convex extension of an element. (C) shows two elements with a type I convexity relation. (D) illustrates a type I.5 relationship, (E) shows a type II relationship, and (F) shows a type III relationship.

It is our goal in this paper only to show that convexity plays a role in amodal completion for which good

[3] An inflection is defined as the point when the sign of curvature changes from positive to negative, and vice versa.

continuation models do not account. So we do not present a full model of convexity-based amodal completion. Specifically, we will consider cases in which convexity favors one completion, while good continuation models either favor the other completion, or make no prediction. We do not claim that convexity makes the correct prediction in all such circumstances, only that it does so in some cases. If true, this will demonstrate that convexity should be incorporated into models of amodal completion, without specifying exactly how it should interact with good continuation cues.

We now present some examples to illustrate the differences between the convexity model and the other two models. There are two key differences. First, the two models based on good continuation do not distinguish between *connecting curves* that are convex, and ones that have a similar shape but are concave. That is, they depend only on the shape of the *connecting curve*, without reference to which side of the curve is figure, and which side is background. An example of this is shown in Fig. 5. On the left, two elements are connected by convex *connecting curves*. On the right, the elements are connected by concave *connecting curves* with the same shape. Both energy and relatability models will not distinguish between these two possible amodal completions, since the shape of possible *connecting curves* will be the same in both. However, the convexity model notes that the elements on the left can be connected convexly, and so are of type I, while the elements on the right have a type II relation. It therefore predicts that the elements on the left will group more strongly together.

A second difference between good continuation and

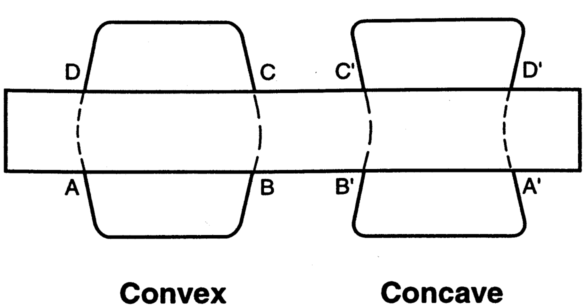


Fig. 5. Schematic illustration of contour completion behind an occluder. In this illustration, the relative distance and orientation between contour segments at A and D are identical to those at A' and D'. We assume that a contour completion behind an occluder depends only on the relative distance and orientation between the two end points, e.g. A and D (in our stimuli the visible contours near the occluder were straight lines). Therefore, a contour completion from A to D is identical to that from A' to D' no matter how one assumes the contour shape should be completed (similarly for contour BC and $B'C'$). Therefore, good continuation models will predict identical strength of amodal completion for the left and right configurations. However, the left one is type I convexity, while the right one is type II. So the convexity model will predict stronger grouping for the configuration on the left.

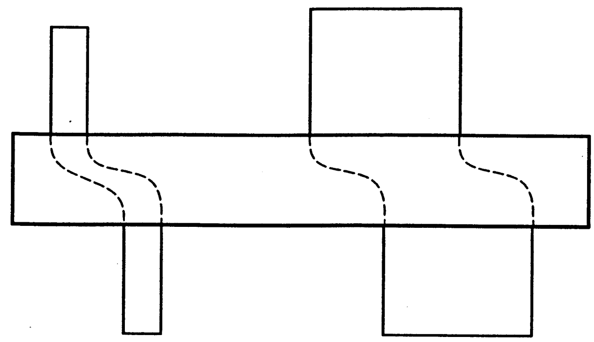


Fig. 6. In this illustration, the grouping on the left has a type III relationship, while the grouping on the right has a type II relationship. However, the relationship between the end points and tangents of the object fragments are identical on both sides. Therefore, all possible *connecting curves* in the two cases are identical, as illustrated by one possible set of *connecting curves*. This example shows a situation in which good continuation models will make no prediction about differential grouping strength, but a convexity model will predict stronger grouping on the right.

convexity is that the prediction of the good continuation models depends only on the shape of the *connecting curves*. The convexity model also takes account of the relationship between the shape of the *connecting curves* and the rest of the element. For example, in Fig. 6, the possible *connecting curves* of the elements on the left are identical to those on the right. Moreover, none of the boundary endpoints are relatable. Unlike the previous example, the possible *connecting curves* on the left and right have the same figure/ground sense. However, the elements on the right have a type II relationship, while the ones on the left are type III. Therefore, convexity predicts a stronger amodal completion on the right, while good continuation models make no prediction.

We will also consider a second approach to applying good continuation models to amodal completion. This second approach argues that we must not only consider how well two elements connect under an occluder into a single entity, we must also consider the strength of the

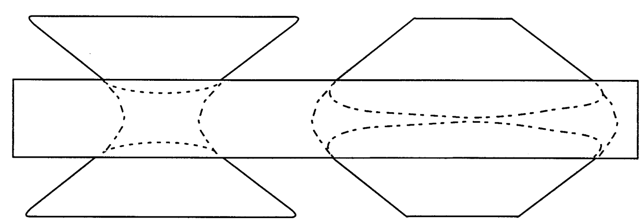


Fig. 7. Schematic illustration of contour completion in the *modified energy* model. The *connecting curves* (the two vertically oriented dashed curves) on the left are identical to those on the right. However, the *closing curves* on the left (the two horizontally oriented dashed curves) are shorter and less curved than those on the right. The model therefore predicts that the connection strength of the figure on the left is weaker. This example illustrates a situation when the figural areas are equalized.

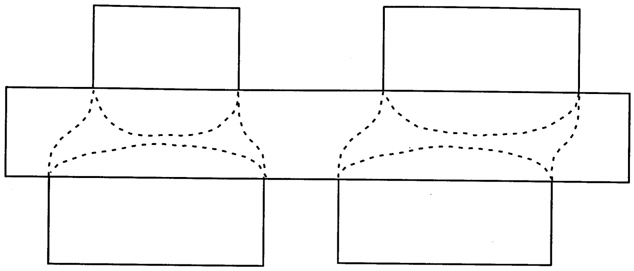


Fig. 8. Schematic illustration of type I.5 (left) versus type II (right) configurations. The *connecting curves* are identical in the two configurations, so are the *closing curves* on the bottom. However, the *closing curve* on the top-right is longer than on the top-left. The *modified energy model* therefore predicts that the one on the right has a stronger grouping. Our convexity model predicts the opposite.

competing hypothesis, that each element completes its boundary to form a separate entity. We will call the possible contours that close an image element without joining it to another *closing curves*. As Fig. 7 shows, there are two competing sets of curves one must consider for these two competing hypotheses, the *connecting curves* and the *closing curves*. In this approach we may expect subjects to group two entities together more strongly when both their *connecting curves* are better according to the energy or relatability model, *and* their *closing curves* are weaker according to these models. Taking this also into account, we create two new models called the *modified energy model* and the *modified relatability model*.

To illustrate the effect of these models, consider Fig. 5 again. The good continuation models made no prediction about these figures, since the *connecting curves* were the same for both the left and right configurations. However, the modified models do not find these configurations equivalent. As shown in Fig. 7, the *closing curves* are quite different on the left and right. The energy model will find the convex configuration (right) to require *closing curves* that are longer and have greater total curvature than those on the left, so they will have a higher energy cost. The relatability model will find that the *closing curves* on the right join edges that are unrelatable, but relatable on the left. So the modified models will predict a preference to see the figure on the left as two separate entities, and will predict stronger amodal completion for the figure on the right. These models now agree with the predictions made by convexity, and we must seek other examples to find a distinctive prediction made by convexity.

To make unambiguous predictions about which of two configurations the modified energy model will prefer, one configuration must have better, or equally good *connecting curves* but weaker or equally good *closing curves*. An example of such a configuration is shown in Fig. 8. The *connecting curves* in the left and right configurations are identical. The *closing curves* in the bottom figures are also identical. However, the *closing curve* on the top element on the right must be longer than the one on the left. The modified energy model predicts that this *closing curve* will be weaker discouraging subjects from seeing this as a separate entity, and encouraging amodal completion. However, the convexity model notes that the elements on the left have a type I.5 relation, while those on the right have a type II relation. This model predicts the opposite of the modified energy model, that the configuration on the left will be preferred as an amodal completion. None of the *connecting* or *closing curves* connect relatable edges, so the modified relatability model makes no prediction here[4].

We have now articulated five possible models of the strength of amodal completion. We have also shown that the convexity model can be distinguished from the other four. It will make predictions that are contrary to those of the other models, or make predictions when the other models do not. If the predictions of the convexity model are true, this will provide evidence that convexity plays a role in amodal completion that cannot be accounted for by existing models that focus solely on good continuation. In the next section we discuss a novel method for judging the strength of amodal completions, which we then use to test these models.

2. Experiments

In the following experiments, we will test the hypothesis that convexity, as opposed to concavity and relatability, is a cue for amodal completion. We hypothesize that this grouping effect will manifest itself in two ways. First of all, grouping will make it more difficult to detect a small stereoscopic depth difference between two planar image regions that are parallel to the image plane and nearly coplanar to each other. Secondly, grouping will bias subjects toward favoring the grouped planar configuration as more likely to be coplanar, even when no depth difference is perceptible.

Our first hypothesis is that perceptual grouping impedes stereoscopic depth discrimination (Liu, Jacobs & Basri, 1995). In other words, the stronger two image patches are grouped together, the more difficult it will be to detect any stereoscopic relative depth between them. Mitchison and Westheimer (1984) have measured discrimination thresholds when subjects decided which

[4] We note that the central axes of the two parts of the figure on the right are more horizontally 'misaligned' than are those on the left. This is generally the case between when we compare parts with a type II relation to those with a type I.5 relation. Therefore, the prediction made by a 'better axis alignment' model would be consistent with the prediction from our convexity model.

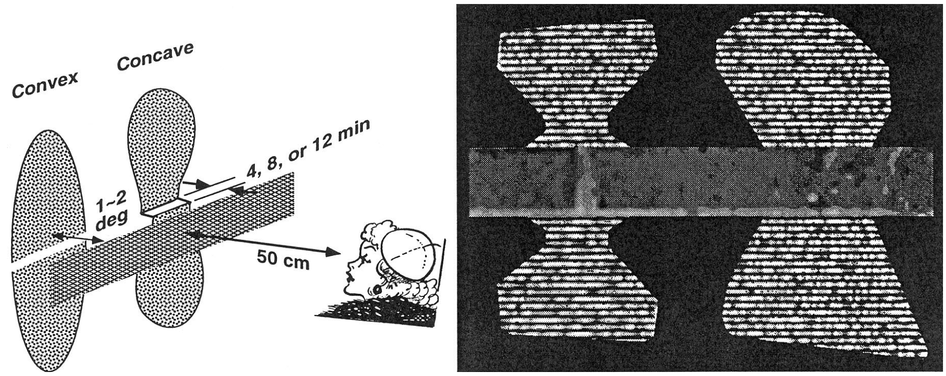


Fig. 9. Left: schematic illustration of the experiment. The subject viewed the stimulus in stereo. Right: example stimulus. The horizontal bar in the middle is the occluder and is closest to the subject. The configuration on the left is called convex, because the simplest completion behind the occluder forms two convex boundary curves. The one on the right is called concave.

of two vertical bars, presented side by side in stereo, was closer in depth. They have found that, when the two vertical bars are connected by two horizontal bars to form into a square, the discrimination threshold increased dramatically. Although Mitchison and Westheimer (1984) used a nearest neighbor model to explain their result, our interpretation is that when the two vertical bars are perceptually grouped (into a square), it becomes much more difficult to discriminate the relative depth between them. Bülthoff, Sinha and Bülthoff (1996) have also found that, when a set of random dots is re-arranged into a human figure, it becomes more difficult to discriminate the relative depth between these dots. They also suggest that two-dimensional form information, that groups the dots into a meaningful configuration, impedes stereo depth discrimination (see also Yin, Kellman & Shipley, 1996). In the following experiments, we will study stereoscopic depth discrimination in amodal completion while manipulating convexity and relatability of contour completions.

Our first experiment demonstrates that it is harder to discriminate relative depth in a convex than in a concave amodal completion configuration. It suggests therefore that convexity is a grouping cue in perceptual grouping. In this case, convexity cues overcome predictions of both the energy and relatability models, indicating that they are not as strong a grouping cue as convexity is. Two subsequent control experiments demonstrate that, once the amodal completion is disabled, the differential effect of discrimination between the two configurations disappears, indicating that amodal completion is responsible for the effect. We then test the hypothesis itself that grouping impedes stereoscopic depth discrimination by comparing depth discrimination between a configuration with amodal completion and the same configuration without. The result suggests that amodal completion makes it more difficult to detect relative depth in stereo.

Next, we study subjects' bias in the same experimental setup when both the convex and concave configurations are coplanar. Subjects show a strong bias in favor of the convex configuration by choosing it most of the time as being coplanar. This suggests that, independent of impeding depth discrimination, amodal grouping induces a bias in the sense that grouped planar regions are perceived as more likely to be coplanar.

Finally, we pit the modified energy model against the convexity model, and demonstrate that the latter better accounts for the data. In this experiment, the modified relatability model makes no predictions, indicating that it too cannot explain our results without taking convexity into account.

2.1. Stimuli

An example stimulus in one experimental trial is shown in Fig. 9. The stimulus was presented in stereo with a viewing distance of 50 cm. The center of the stimulus was a brick-patterned horizontal occluder, which was in front of the monitor plane. The width of the occluder was 30 cm, with the height varying from 4 to 8 cm. The four textured polygons behind the occluder were in two depth planes. Three of them were behind the monitor plane in the range of 0 to 2.5 min of arc[5]. The fourth was either in front of or behind it in one of three depths: 4, 8 or 12 min of arc. The stimulus was presented with orthographic projection, so that the statistics of the texture pattern on the four polygons were identical, thereby providing no relative depth cues. The texture itself was created by interleaving horizontal stripes of random 'checkers' and black horizontal stripes. Each 'checker' was about $2.7' \times 2.7'$ arc in visual angle.

[5] We subsequently verified that it was not critical for the three polygons to be coplanar. When the third was at a different depth plane, similar results were obtained.

The two polygons on the left side in Fig. 9 amodally complete themselves behind the occluder to form a two-dimensional convex region there, we call the completed figure the convex configuration. Similarly, the two on the right form a concave region, we call it the concave configuration. We deliberately created concavities along the visible contours of the convex configuration, so that subjects would see both convexity and concavity and not be biased toward seeing only convex contours. We also created asymmetric configurations so that symmetry would not play a role in our experiments.

There are four T-junctions on the left side and four on the right. The relative distances between the four T-junctions on the left are identical to those on the right (the absolute positions of the four T-junctions on the right side could be flipped vertically and horizontally). Therefore the relative distance between the two endpoints of a *connecting curve* on the left side is the same as on the right. In addition, the stem of any T-junction is locally a straight line, therefore the endpoint curvature of any completion curve is zero and plays no role in the energy formulation. There are two conditions regarding the relative tangent orientations of each *connecting curves*, for the convex and concave configurations, respectively, as described below.

2.1.1. Condition 1: equal completion energy

The relative tangent orientation of a *connecting curve* of the left figure is the same as that on the right, except one is convex and the other concave. The absolute tangent direction at each T-junction is sampled from a uniform distribution of [20°, 70°) (relative to the horizontal).

Therefore, the distribution of the relative tangent orientation is the convolution of two such uniform distributions. Each contour completion can be either relatable or non-relatable. The energy and relatability models will make no predictions regarding the depth discriminabilitv of the convex and concave configurations.

2.1.2. Condition 2: non-relatable convex versus relatable concave

For the convex configuration, the tangent orientation at each T-junction is drawn from a uniform distribution [20°, 50°). For the concave configuration, it is [40°, 70°) for one *connecting curve* and [45°, 75°) for the other. Therefore the convex configurations are almost always non-relatable, whereas the concave are almost always relatable. The relatability theory (Kellman & Shipley, 1991) will predict a stronger completion for the concave than for the convex configurations. The small number of cases when the convex completion is relatable and, independently, the concave completion is non-relatable will serve as a test of the extent to which 90° is critical for the relatability theory.

2.2. Apparatus and experimental procedure

All stimuli were displayed on the monitor of a Silicon Graphics computer. Subjects wore shutter glasses (StereoGraphics Co., CA) that synchronized with alternating monitor frames in order to view the stimuli in stereo.

In each trial, the convex and concave configurations were presented side by side for unlimited time until the subject responded. Randomly, one configuration was coplanar, the other had a depth difference. The subject decided whether the left or the right configuration was non-coplanar by pressing the corresponding mouse button. No feedback was provided. The left–right positions of the two configurations were randomized, so was the upright orientation of each configuration. After the subject responded, the next trial started automatically. The two experimental conditions were randomly interleaved. Each subject ran 600 trials, which lasted for about 45 min.

Five naive subjects and author ZL participated in the experiment.

2.3. Results

The subjects' accuracy of stereo depth detection was analyzed as a function of contour completion (convex vs concave), the three depth steps, and the two experimental conditions. ANOVA revealed a significant difference between the convex and concave configurations (Fig. 10, $F[1,5] = 13.97$, $P < 0.01$): it was easier to detect a stereo step for a concave than for a convex configuration (87.15 vs 72.08%). This suggests that a convex contour completion is grouped more strongly than a concave one. As expected, different depth values gave rise to different detection accuracies ($F[2,10] = 265.18$, $P < 0.001$), the larger the depth difference, the easier the detection was (91.35, 82.83 and 64.66%). Interestingly, the two experimental conditions (first: convex vs concave with equal

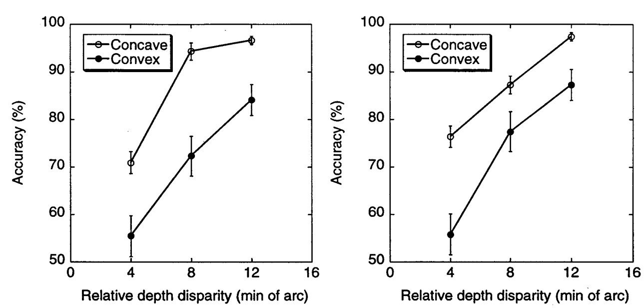


Fig. 10. The subjects' accuracy in discriminating which configuration was non-coplanar. It was always harder to detect a depth step for a convex configuration. Left: when the completion contours were identical for both configurations except their convexity. Right: when the completion contours always curved more (> 90°) for the convex than for the concave configuration (< 90°).

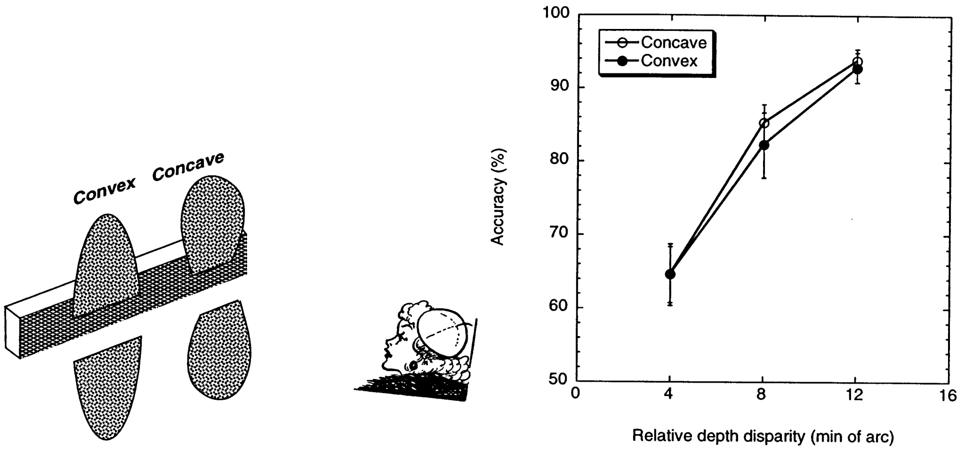


Fig. 11. Left: schematic illustration of the first control experiment. The 'occluder' was furthest away from the subject, so no perceptual grouping was possible. Right: eight subjects' performance. No difference was found between the 'convex' and 'concave' configurations.

completion energy; and second: non-relatable convex vs relatable concave) did not yield a significant difference ($F[1,5]=2.31$, $P=0.2$). This suggests that relatability plays an insignificant role in the current experiment.

We further analyzed the situation when convex completions were strictly non-relatable, while the concave were always relatable. Again, convex configurations were more difficult than the concave for depth detection (51.00 vs 75.70%, $F[1,5]=8.50$, $P<0.03$). This result contradicts the predictions of both the *energy* model and the relatability theory.

In addition, we analyzed (while collapsing the two conditions) the cases when both the convex and concave completions were relatable, and when both were non-relatable. Again, relatability does not appear to play a significant role. Relatable: convex 39.50%, concave 78.22%, $F[1,5]=8.66$, $P<0.03$. Non-relatable: convex 42.94%, concave 80.94%, $F[1,5]=13.67$, $P<0.01$. These results suggest that relatability does not account well for amodal completion. Rather, the results concurred with the hypothesis that convex completion is a grouping cue.

2.4. Control experiments: is amodal completion critical?

2.4.1. 'Occluder' behind the figures

In order to test whether the above results are indeed due to perceptual completions behind an occluder as opposed to some other stimulus factors such as visible areas, we conducted a control experiment in which perceptual completion is disabled while everything else is the same. This was done by using stereo to move the 'occluder' behind the four polygons, so that there were visible gaps between the polygons that would have completed otherwise (Fig. 11). The visible surface area was a concern because the visible areas of the two textured polygons on the left did not precisely match those on the right. Hence a larger visible area may give rise to a stronger stereo disparity signal (but see Richards & Kaye, 1974).

Eight subjects participated in the experiment. As shown in Fig. 11, when no perceptual completion was possible, there was no difference between the 'convex' and 'concave' configurations. This suggests that completion behind an occluder is important for the differential result in the previous experiment.

2.4.2. A shrunken 'occluder'

In the control experiment above, when the amodal completion was disabled, the role of foreground-background of the 'occluder' and the four polygons was also switched. To check whether this was responsible for the results above, we conducted a second control experiment in which we shrank the height of the 'occluder' to 70% of its original height. Therefore, there was a visible gap between each polygon and the 'occluder', regardless of whether the 'occluder' was in front of or behind the polygons.

Eight subjects participated in this experiment. As shown in Fig. 12, no difference was found between having the 'occluder' in front or behind, and between the 'convex' and 'concave' configurations. This further suggests that amodal completion is responsible for the differential results between the convex and concave configurations.

2.5. Testing the hypothesis

One central hypothesis in our study is that perceptual grouping impedes stereoscopic depth discrimination. We conducted an experiment to directly test this hypothesis. The stimuli differed from those above as follows. The polygons on the left were identical to those on the right (Fig. 13). They were both convex or both concave. Yet, one configuration was perceptually completed behind the occluder, the other was not. The latter was achieved by shrinking the height of its 'occluder' by 30%, and the 'occluder' was either in the front or behind.

Three subjects participated in the experiment. As shown in Fig. 13, when the depth step was sufficiently small, it was harder to detect this depth step for the completed than for the uncompleted ($F[2,4] = 1.48$, $P < 0.05$). Another significant, and expected, effect is that the larger the depth step, the easier the detection is ($F[2,4] = 29.90$, $P < 0.004$). This result supports the hypothesis that perceptual grouping impedes stereoscopic depth detection.

2.6. Decision bias and learning

Our working hypothesis has been that subjects will perceive fragments that form a better unit as more likely to be coplanar. One implication of this hypothesis is that in our experimental setup there may be a decision bias for subjects to select convex stimuli as coplanar. That is, even when no perceptible depth difference exists for either the convex or the concave stimuli, subjects will be inclined to choose the convex as more coplanar. Such a decision bias would have several implications, some of which could have effected the results in the experiments we have described so far. We will discuss these implications briefly, and describe a new test, specifically aimed at detecting this decision bias.

If subjects have a bias to perceive the convex stimuli as coplanar, regardless of stereo cues, then they will more often choose the concave stimuli as being noncoplanar. As we report the accuracy rate for concave stimuli, it indicates, given all trials in which the concave configuration were not coplanar, the fraction of times the subject chose the concave configuration. Therefore, a decision bias would result in better performance when the correct answer is concave, and poorer performance when it is convex. For example, a subject who randomly selected the concave configuration as the correct answer, regardless of stereo cues, would achieve 75% performance for when a disparity in depth was present for concave stimuli, and 25% for convex. Therefore, our results showing better performance for concave than convex configurations may be partly due to a decision bias. If this is the case, disabling amodal completion by shrinking the occluder or moving it behind the stimuli may result, not only in improved performance on the convex stimuli, but might result in decreased performance for concave stimuli, by reducing the decision bias. However, the absence of amodal completion might be a factor for improving performance under all conditions. Hence, while we can correctly predict that disabling amodal completion should improve performance for convex stimuli both by removing grouping and indirectly reducing the decision bias, no prediction is possible on the effect of disabling amodal completion on overall performance for concave stimuli.

So far, our results are consistent with the possibility of a decision bias. We now test this possibility directly. We show that a decision bias is present, although it does not account for all our results. One alternative to our current experimental method is to present only one stimulus configuration, convex or concave, behind the occluder. A depth step would always exist, with either the upper or the lower part in front of the other. The subject would then decide whether the upper patch is in front of or behind the lower one. The advantage of this alternative is that it gives rise to a bias-free measure of depth discrimination. However, this would also result

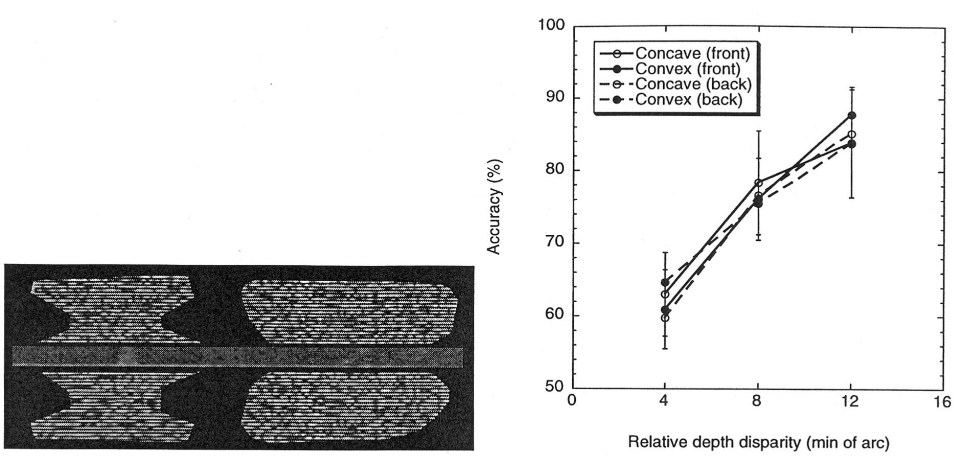
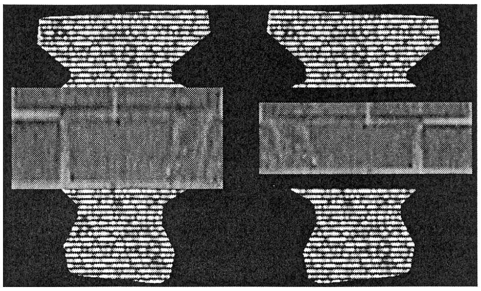


Fig. 12. Left: example stimulus in the second control experiment, where a gap is clearly visible between the shrunken 'occluder' and the four polygons. Right: no difference was found between the 'occluder' in front and behind conditions, and between the 'convex' and 'concave' configurations.



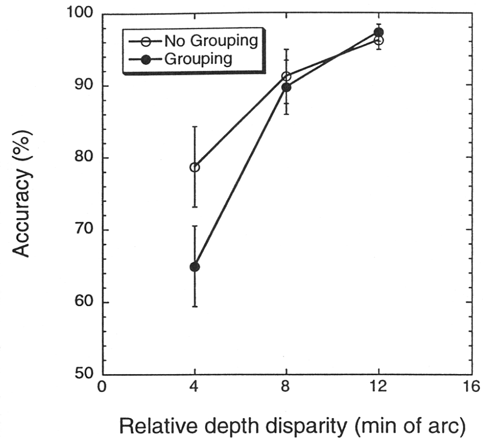


Fig. 13. Left: an example stimulus when the irregular shaped figures on the left are identical to those on the right (with a left–right reversal). The occluder on the left was closest in depth to the subject, therefore the left configuration amodally completes behind the occluder. The height of the 'occluder' on the right is shrunken, therefore no perceptual completion was possible, regardless of whether it is in the front or behind. Right: it is more difficult for the subjects to detect the depth step for the amodally completed configuration.

in a much more indirect comparison between a convex and a concave configuration that have identical contour completions. In order to keep the direct comparison and, at the same time, separate bias from discrimination sensitivity, we repeated the first experiment with two fresh naive subjects and added trials when both configurations were coplanar. Their result is shown in Fig. 14. Apparently, both subjects had a strong bias to perceive the convex configuration as more likely to be coplanar when both configurations were actually coplanar.

Having confirmed that there is indeed a bias favoring convex configurations as more likely to be coplanar, we tested four more subjects who had participated in the experiments in this study and therefore were well practiced. As shown in Fig. 14, these experienced subjects have learned to reach a bias free judgment (when both configurations are coplanar), but it was still more difficult to detect the depth step for the convex than for the concave configurations ($F[1,3] = 11.61$, $P < 0.04$). These results suggest that in addition to a bias that favors convex configurations as more likely to be coplanar, convex configurations indeed impede stereo depth discrimination.

2.7. Testing the modified energy model

Thus far, we have used convex vs. concave configurations to test the *energy* model when the 'cost' of completing the contours with *connecting curves* is identical for both the convex and concave configurations, or is even lower for the concave configurations. In this case, the *energy* model predicts no preference between the two configurations, or predicts a preference for the concave configuration, whereas human subjects perceive the convex configuration as having a stronger grouping. However, as we discussed earlier, a *modified energy model* would argue that the grouping strength of any configuration depends on how strongly two parts are grouped together relative to how strongly they can be separated. In other words, it is more sensible to consider the overall grouping strength as the 'cost' of grouping two parts with the *connecting curves* relative to the 'cost' of separating them with *closing curves*.

In this experiment, we will use two configurations that keep constant the 'cost' of the *connecting curves* but differ in the *closing curves*. We will demonstrate that human performance contradicts the prediction of the *modified energy model* but concurs with the convexity model.

2.7.1. Stimuli

Fig. 15 shows an example stimulus in this experiment. The *connecting curves* are identical for the left configurations (type II) and for the right (type I.5). The *closing curves*, however, are different. It was harder, as predicted by the modified energy model, to self-close each of the two parts on the left than on the right side. This is true because the longer gap requires a longer, and more costly *closing curve*. Therefore, the *modified energy model* would predict that the type II configuration (left) groups more strongly than the type I.5 (right). The stereoscopic depth differences used in this experiment were 3, 6, 9 and 11 min of arc in visual angle.

Six subjects, who were unaware of the experimental purpose and had no psychophysics experience, plus author DWJ, participated.

2.7.2. Results

It turns out that, when the depth step was relatively large (9 and 11 min of arc), no difference was found. However, when the depth step was small (3 and 6 min of arc), it was easier to detect the relative depth differ-

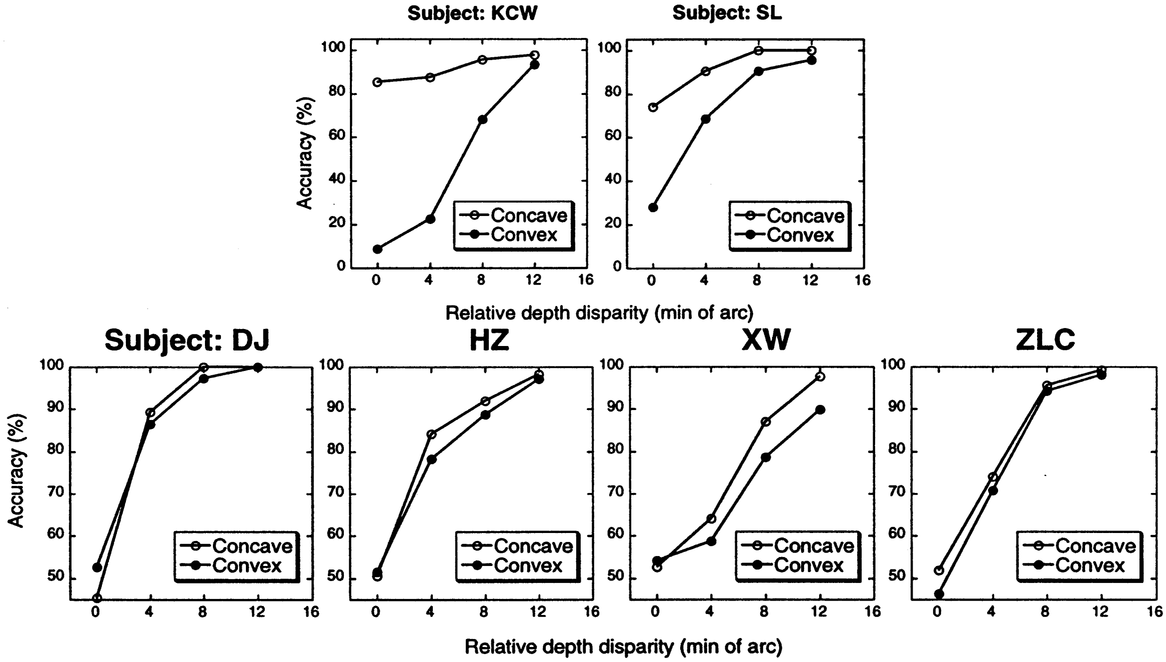


Fig. 14. Top: results from two novice subjects who showed substantial bias in perceiving the convex completion as more likely to be coplanar (the zero disparity in the horizontal axis represents the condition when both the convex and concave configurations were coplanar). Bottom: results from four experienced subjects whose overall performance was improving and whose performance difference between the convex and concave conditions is being reduced through extensive practice. Importantly, no bias existed for these subjects, but it was still more difficult to detect the depth step for the convex than for the concave configurations.

ence for the type II stimuli than for the type I.5 stimuli (83.0 vs 79.2%, $t(6) = 2.60$, $P < 0.02$). This suggests that the type I.5 stimuli were grouped more strongly than the type II ones.

3. Discussion

Our paper has addressed two main points. First, we have demonstrated that stronger perceptual grouping causes a tendency in subjects to perceive planar figural elements as coplanar. This tendency can be measured by the extent to which it impedes the ability of subjects to detect non-coplanarity using stereo cues. This result suggests that perceptual grouping cues can override even low-level visual cues about the nature of scene structure. In doing so, we also provide an alternate explanation for the closely related prior results of Mitchison and Westheimer (1984). While this result may be of independent interest, we have primarily used it as an experimental tool to probe the role of convexity in perceptual grouping. The role that perceptual grouping plays in inhibiting coplanarity detection may also provide a useful methodology for exploring other aspects of perceptual grouping, since it provides us with a method of judging the quantitative effect of competing perceptual grouping cues.

Second, we have shown that convexity plays a role in perceptual grouping that cannot be accounted for by existing models based on good continuation. Models based on curves of least energy do not predict convexity effects at all. Inflections, for example, which signal the transition from a convex to a concave part of a contour, do not necessarily result in higher energy curves. This is because an inflection point has zero curvature, so surrounding points can all have low curvature. To some extent the relatability model accounted for this by penalizing inflections. However, the presence or absence of inflections are not sufficient to judge the convexity or concavity of a grouping. As we show in many examples, connections with the same numbers of inflections

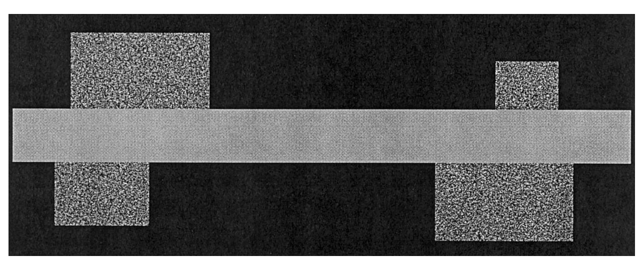


Fig. 15. An example of a type II (left) versus type I.5 (right) stimulus, presented in stereo in the experiment. The 'cost' of 'energy' is the same to connect the two textured parts behind the horizontal occluder on the left and on the right side. The 'cost' of 'energy' to separate the two textured parts on the left, however, is higher than on the right, as predicted by any *good continuation* contour completion model. Therefore, when both costs are considered, the *modified energy model* predicts that the type II configuration (left) has a stronger grouping. Our experimental result suggests the opposite.

can signal very different types of convexity relationships. When this occurs, we have shown that the type of convexity relationship is the key factor in determining the strength of perceptual grouping. In fact, when convexity signals a different organization than does low energy or relatability, we have shown that convexity can be the stronger cue.

Although we have argued for the importance of convexity in grouping, we have not presented a precise model of grouping. For example, we do not argue that only convexity relationships are important in determining the strength of grouping. There is much evidence (e.g. Kellman & Shipley, 1991; Field, Hayes & Hess, 1993) that good continuation plays an important role in perceptual organization, and that contours that turn small angles lead to stronger groupings than those that must be connected at acute angles. Consequently we feel that convexity should augment, not supersede good continuation models.

Another hypothesis might be that convexity type is always a stronger grouping cue than good continuation. This would imply, for example, that any figural elements with a type I relationship will always group more strongly than figural elements with a type II relationship, even if the type II configuration has a better connection according to good continuation models (assuming that other factors, such as distance, are held constant). In this view, good continuation would only generate preferences between potential groups that have identical convexity relationships. We have shown one example in which this is the case in Section 2.3. However, it is also consistent with our data to suppose that convexity and good continuation provide signals to perceptual grouping that are combined, so that a strong cue from good continuation can override convexity cues. We leave it as an open question whether this does occur.

Finally, we wish to point out that our experiments are relevant not only to energy and relatability models, but to a whole class of models of which these are just examples. We have defined the *connecting curves* as the curves needed to join two figural fragments, and *closing curves* as the curves needed to separate them. Energy and relatability, and the modified versions of them that we have considered, are examples of models that say that the strength of a grouping will depend only on the possible shapes of the *connecting* and *closing curves*. However, in Fig. 15 we have shown an example in which *closing* and *connecting curves* are identical, except for the length of one pair of *closing curves*. If we assume that longer distances signal weaker perceptual completions, then any model based solely on analyzing the *closing* and *connecting curves* must make predictions that are contrary to our experimental results. To deal with stimuli such as these, models must be enriched by taking account of the relationship between object contour and the interior regions of objects. This allows us, for example, to incorporate convexity relationships fully into a model of perceptual grouping.

Acknowledgements

Part of this work was done when ZL was a visiting scientist at the NEC Research Institute (NECI), Princeton, NJ. We thank Patrick Cavanagh, Robert de Ruyter van Steveninck, James Elder, Pascal Mamassian, John Oliensis, and Laurie Wilcox for helpful discussions. James Elder pointed out to us the work of Mitchison and Westheimer (1984). The research of Ronen Basri was supported in part by the United States–Israel Binational Science Foundation, grant No. 94-00100. Ronen Basri is an incumbent of Arye Dissentshik Career Development Chair at the Weizmann Institute of Science.

References

Bertamini, M., & Friedenberg, J. D. (1999). Effects of contour convexity on performance in symmetry detection and in relative position judgments. In *ARVO Abstracts. Investigative Ophthalmology & Visual Science*, vol. 40.

Brady, M., & Grimson, W. E. L. (1981). The perception of subjective contours. Technical report, Massachusetts Institute of Technology.

Bülthoff, I., Sinha, I., & Bülthoff, H. H. (1996). Top-down influence of recognition on stereoscopic depth perception. In *ARVO Abstracts. Investigative Ophthalmology & Visual Science*, vol. 37 (5168 p.).

Driver, J., & Baylis, G. C. (1995). One-sided edge assignment in vision: 2. Part decomposition, shape description, and attention to objects. *Current Directions in Psychological Science*, *4*, 201–206.

Field, D., Hayes, & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local 'association field'. *Vision Research*, *33*, 173–193.

Gibson, B. S. (1994). Visual attention and objects: one versus two or convex versus concave? *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 203–207.

Hoffman, D., & Richards, W. (1984). *Parts of recognition* (pp. 76–82). Cambridge, MA: MIT Press.

Horn, B. K. P. (1981). The curve of least energy. Technical report, Massachusetts Institute of Technology.

Jacobs, D. W. (1996). Robust and efficient detection of salient convex groups. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *18*, 23–37.

Kanizsa, G., & Gerbino, W. (1976). Convexity and symmetry in figure-ground organization. In *Vision and artifact* (pp. 25–32). New York: Springer.

Kanizsa, G. (1979). *Organization in vision*. New York: Praeger.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.

Liu, Z., Jacobs, D. W., & Basri, R. (1995). Perceptual completion: beyond good continuation. In *ARVO Abstracts. Investigative Ophthalmology and Visual Science*, vol. 36 (pp. 2219).

Mitchison, G. J., & Westheimer, A. (1984). The perception of depth in simple figures. *Vision Research*, *24*, 1063–1073.

Mumford, D. (1994). Elastica and computer vision. In *Algebraic geometry and its applications*. New York: Springer.

Nitzburg, M., & Mumford, D. (1990). The 2.1-D sketch. In *International Conference on Computer Vision* (pp. 138–144).

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

Richards, W., & Kaye, M. G. (1974). Local versus global stereopsis: two mechanisms? *Vision Research*, *14*, 1345–1347.

Rock, I., & Palmer, S. (1990) The legacy of Gestalt psychology. *Scientific American*, July, 84–90.

Rubin, E. (1927). Visuell wahrgenommene wirkliche bewegungen. *Zeitschr*. *Psychologie*, *103*, 354–384.

Shashua, A., & Ullman, S. (1988). Structural saliency: the detection of globally salient structures using a locally connected network. In *2nd International Conference on Computer Vision*, Clearwater, FL (pp. 321–327).

Trytten, D., & Tuceryan, M. (1991). Segmentation and grouping of object boundaries using energy minimization. In *IEEE Conference Computer Vision and Pattern Recognition* (pp. 730–731).

Ullman, S. (1976). Filling-in the gaps: the shape of subjective contours and a model for their generation. *Biological Cybernetics*, *25*, 1–6.

Williams, L. R., & Hanson, A. R. (1996). Perceptual completion of occluded surfaces. *Computer Vision and Image Understanding*, *64*, 1–20.

Williams, L. R., & Jacobs, D. W. (1997). Stochastic completion fields: a neural model of illusory contour shape and salience. *Neural Computation*, *9*, 837–858.

Yin, C., Kellman, P. J., & Shipley, T. F. (1996). Surface completion influences depth discrimination. In *ARVO Abstracts*. *Investigative Ophthalmology & Visual Science*, vol. 37 (1324 p.).