PERGAMON

AVAILABLE AT www.ComputerScienceWeb.com POWERED BY SCIENCE DIRECT®

Neural Networks 16 (2003) 915–923

Neural Networks

www.elsevier.com/locate/neunet

2003 Special issue

Interpolation processes in the visual perception of objects

P.J. Kellman

University of California, 405 Hilgard Avenue, Los Angeles, CA 90095-1563, USA

Abstract

Visual perception of objects depends on segmentation and grouping processes that act on fragmentary input. This paper gives a brief overview of these processes. A simple geometry accounting for contour interpolation is described, and its applications to 2D, 3D, and spatiotemporal object interpolation processes are considered. A method is described for distinguishing interpolation based on this geometry from more global or top-down influences. Results suggest a separation between interpolation based on relatively local stimulus relations, which give rise to precise boundary representations, and processes involving recognition from partial information, which do not. Aspects of the model—especially the unified treatment of illusory and occluded objects—raise questions about the nature of seeing. Although it is often believed that illusory objects are *perceived*, while occluded objects are *inferred*, I suggest that both research and theory converge in supporting a more unified account. Illusory and occluded contours and surfaces do not divide into the real, the perceived, and the inferred, but are all *represented*, and in key respects, derive from identical perceptual processes.

Keywords: Interpolation processes; Visual perception; Objects; Occlusion; Contours; Illusory contours

1. Introduction

Few verbs have the multiplicity of meanings as 'to see.' One might see a point made in discussion, see that it's getting late, or see about dinner reservations. In understanding visual perception, one of the most important ambiguities involves what it is that we can see. A long and successful scientific tradition has characterized the eye as an optical system and described visual pathways as allowing us to see light or extract properties of light. Conversely, in common parlance and in the understanding of perception, we seem to be involved with the seeing of *objects*, as well as their arrangement in space and events involving them. Do we see light? Or objects? Assuming that the answers are 'yes' in some sense to both questions, what is the relation between these different kinds of seeing?

These questions may seem to be purely philosophical, but they become crucial in the study of perceptual organization. In this article, we consider specifically perceptual processes that create objects from what is often a fragmentary patchwork of light reflected to the eyes from a given object in the world. I will both describe some current directions in research on object perception and consider a couple of larger issues involving the dichotomy between receiving reflected light and achieving perceptual representations.

The focus will be on processes of perceptual organization that build representations of complete contours, objects, and surfaces from fragmentary input. To keep the discussion brief, I provide in most cases only the gist of research findings, referring to other published work for details.

Perceptual systems provide the information that humans and other organisms use to guide thought and action. Vision is special among these systems, in that it gives detailed knowledge of objects and the spatial layout, and it works at a distance. Understanding vision requires us to determine how visual information leads to descriptions of objects, spatial arrangements, and events.

Our descriptions of perceived reality are heavily based in a vocabulary of *objects*—bounded volumes of matter that function as units in our interactions with the world. Objects are both physical and psychological units; in fact, they may be best described as ecological units (Kellman & Arterberry, 1998). In physical reality, we believe that there exist coherent entities, such as pebbles, books, and cars. They separate from other things and cohere over movement. At the same time, our psychological notions of objects are restricted in terms of size, duration, and degree of cohesion. Neither the planet earth nor a sub-atomic particle is an object for ordinary cognition and behavior. So, objects are physical units at scales relevant to our thought and behavior.

How do we perceive objects? Although the operation of object perception is ordinarily quick and effortless,

0893-6080/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0893-6080(03)00101-1

the scientific explanation of object perception is extremely challenging. Many of the problems are commonly considered together as issues of *visual segmentation and grouping*. Whether the problem is segmentation or grouping depends on the description with which you start. If the starting point is an image of a scene, either on your retina or on a CRT screen, the issue is how to segment this entity into parts that correspond to different objects in the world. If instead you start with the well-known coding of visual inputs in mammalian vision into small, local patches of oriented contrast in early cortical areas, the problem is how to group together the outputs of thousands of units that might all be registering parts of a single object. Grouping is also the issue if one starts with an image description in terms of many pixels. Likewise, if we start with visible regions, parts of a single object must be grouped, as often they project to several different retinal locations, as we will consider.

A pervasive and fundamental difficulty in achieving descriptions of objects from information in reflected light may be labeled the *fragmentation problem*. Whereas the world and our representations of it contain coherent objects and continuous surfaces, the input from the world to our eyes is fragmentary. Most objects are partly occluded; their projections to the eyes are interrupted by parts of other objects. Viewing a scene through foliage, the observer's retina receives many separate patches of each object in the scene. How do we obtain representations of complete objects—a building, a car, a person—under these circumstances? The answers are important, as these circumstances are more the rule than the exception in ordinary perception. The problems become all the more complicated when observers or objects move, as patterns of occlusion continuously change.

2. Visual interpolation phenomena

Fortunately the visual system has processes that overcome fragmentation in the input and produce representations of complete objects. Shown in Fig. 1 are examples of several phenomena in which the visual system overcomes gaps in the input in perceiving objects. Fig. 1A shows a case of partial occlusion; the three black regions all appear as part of a complete object, whose shape is apparent. Fig. 1B shows an example of the phenomenon of *illusory contours* or illusory objects. Here, contours are connected across gaps such that these interpolated parts appear in front of other surfaces. In Fig. 1C, an additional effect, one of transparency, can be observed in the figure that is created across the gaps.

These phenomena are formally similar in that equivalent or similar visible contours and surface fragments become linked to form objects, despite intervening gaps. Research suggests that the similarity is more than formal: these phenomena share a common underlying interpolation process.

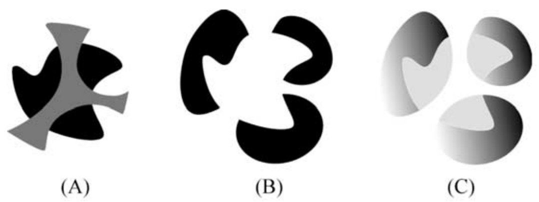

Fig. 1. Visual interpolation phenomena: (A) partial occlusion; (B) illusory contours; (C) transparency.

3. Visual interpolation processes

How do visual processes overcome fragmentation to achieve accurate object perception? The visual system uses relationships among visible *contour* segments to guide segmentation and derive shape. It also uses relationships of *surface* properties to establish surface continuity. Most of our focus will be on contour processes, but we will describe surface processes briefly in a later section. After giving a sketch of these processes, I will consider their implications for the nature of seeing: do these processes create illusions or give us windows into reality?

3.1. Contour extraction processes

Object perception involves a number of conceptually distinct information processing tasks and representations. (These may or may not be realized as separate processes or mechanisms in the brain.) Fig. 2 provides a schematic model of these processes, representations and their interactions. Some components and their connections reflect established findings whereas others represent newer conjectures. Considerable differences exist in how much is known about precise computations and neural implementations in the specific boxes. Rectangles indicate functions or processes, and octagons indicate representations. The model has few representations: output representations of shape and unity (labeling regions belonging to a single object) and one intermediate representation: the *visible regions representation*. Among other things, this representation marks connected visual areas, derived from depth, velocity, and luminance maps, and labels their contours as bounding or belonging to other (occluding) objects. (For this and other details of the model, see Kellman, Guttman, & Wickens, 2001).

Our immediate concern is the contour stream, which locates discontinuities in luminance, color, texture, depth, and motion, uses these to find edges and junctions, and ultimately leads to meaningful contours and object shape. The importance of contours derives from the fact that objects will tend to be relatively homogenous in their composition (and thus in their lightness, color, texture,

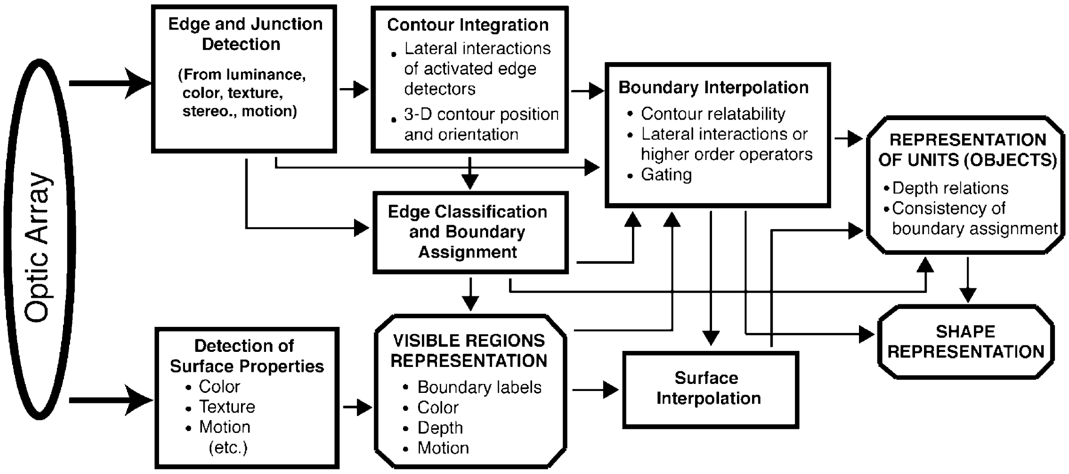

Fig. 2. A framework for object perception. Rectangles are functions; octagons are representations (see text).

depth, etc.) and will often differ from nearby objects and surfaces. Visible contours often, but not always, mark the locations of object boundaries.

Computations leading to perception of contours and edges appear to begin in the early cortical areas V1 and V2. Cells in these areas respond to oriented luminance contrast at particular spatial frequencies in particular retinal locations (Campbell, Cooper, & Enroth-Cugell, 1969; Hubel & Wiesel, 1968). By area V2, and perhaps earlier, many cells respond selectively to particular binocular disparities, providing the basis for stereoscopic depth perception (Fischer & Poggio, 1979). Some cells in the early cortical areas also respond preferentially to motion, although areas upstream, particularly area V5 (the human homologue to macaque area MT), appear to be specialized for motion processing. Numerous proposals in both biological and computational vision have been made about how early cortical responses can be utilized to detect luminance edges. The modeling of object perception based on other types of edge inputs, such as discontinuities in stereoscopic depth and motion is less advanced. Likewise, how the visual system integrates edge information arising from various sources remains relatively unknown.

Additional tasks that we will not discuss here include the classification of edges, as arising from illumination or reflectance differences, and boundary assignment of those edges determined to be occluding edges.

3.2. Contour interpolation

Although the tasks of contour finding, classification, and boundary assignment are challenging in themselves, even their complete solution does not provide reasonable representations of objects. Because of occlusion, in normal scenes, most objects are partly occluded by others. *Contour interpolation* processes connect visible contours across gaps caused by occlusion to produce perceptual units that correspond more accurately to the actual objects in a scene.

3.2.1. Tangent discontinuities and contour relatability

In recent years, we have learned a great deal about how interpolation processes proceed (for more details, see Kellman et al., 2001; Kellman & Shipley, 1991). These computations begin with contour intersections or *junctions*. Junctions are points of *tangent discontinuity*—places where contours have no unique orientation. Most often there are two orientations at image junctions. As can be seen in Fig. 1, interpolated contours begin and end at these tangent discontinuities in the image. These points include all of the places where edges go out of sight and need to be interpolated. It has been proposed that “end-stopped” cells in visual cortex (cells sensitive to the contours that end in their receptive fields) underlie junction detection (e.g., Heitger et al., 1992).

Interpolated contours not only begin at tangent discontinuities, but they agree with the slopes of visible contours that terminate at these connection points (see examples of Fig. 1). Contour interpolation can be summarized by saying that visual system interpolates according to a smoothness principle, called *contour relatability*, such that interpolated contours are smooth (differentiable at least once) and monotonic (singly inflected), and at their endpoints, they match the slopes of real contours. Fig. 3 shows a geometric construction defining relatability (Kellman & Shipley, 1991). Except for the additional constraint that interpolated edges bend by no more than 90°, relatability can be summarized by saying that edges are relatable if their linear extensions meet in their extended regions.

The notion of relatability in perceptual grouping is not the same as but is related to the principle of *good continuation*, proposed as one of several principles of perceptual organization in the early 20th century by the Gestalt psychologists (Wertheimer, 1958). Relatability

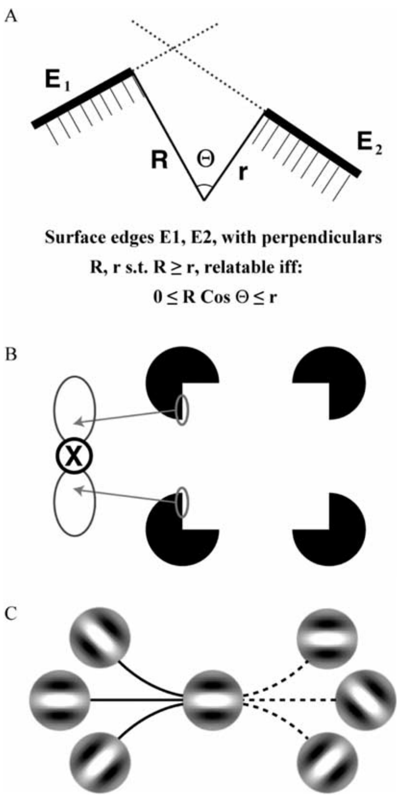

Fig. 3. (A) Relatability geometry (Kellman & Shipley, 1991); (B) bipole implementation of interpolation (Grossberg & Mingolla, 1985; Heitger et al., 1998); (C) network implementation (Field et al., 1993).

appears to tap into important regularities about the smoothness and predictability of objects (Geisler, Perry, Super, & Gallogly, 2001).

3.2.2. Quantitative variation in interpolation

The definition of relatability as given is categorical—it distinguishes contours that can be connected by interpolation from those that cannot. There is an important categorical aspect to these perceptual decisions, as perhaps their most important function is to allow us to perceive visible regions as connected or not. Even so, experimental data suggest that within the category of relatable edges, there are quantitative variations in strength (Field, Hayes, & Hess, 1993; Kellman & Shipley, 1991). Interpolation strength decreases as the angular relation between edges goes from collinear to near-zero at about 90°. It also rapidly declines with misalignment of parallel edges (Shipley & Kellman, 1992a).

Experimental results suggest that strength of contour interpolation depends not only on angular relations of edges but on their lengths and separation (Banton & Levi, 1992; Shipley & Kellman, 1992b). Specifically, over a wide range of image sizes, strength of interpolation seems to be proportional to: $(L_1 + L_2)/(L_1 + L_2 + g)$, where L_1 and L_2 are the lengths of (physically specified) relatable edges and g is the gap between them (Shipley & Kellman, 1992b).

3.2.3. Implementation of the relatability geometry

How are the relationships captured by the geometry of relatability implemented in neural mechanisms of vision? This is an active area of research. Several neural-style models of contour interpolation have been proposed (Grossberg & Mingolla, 1985; Heitger, von der Heydt, Peterhans, Rosenthaler, & K bler, 1998). There are two basic ideas about the mechanisms used by the visual system to surmount gaps. One is the use of higher-order operators, often called bipole cells. Real edge inputs on either side of a gap feed into these nonlinear grouping operators (Heitger et al., 1998). The activation of these operators, centered over a discontinuity in edge input, may be used to construct a contour that spans the gap (Fig. 3B).

Interpolation may also be carried out by an interactive network of orientation-signaling units (Field et al., 1993; Fig. 3C). According to this idea, luminance-defined edges activate some units, leading to facilitative interactions with other units that do not receive direct stimulus input. Interpolation occurs when a path of units, defined on either end by directly stimulated units, becomes active as the result of these interactions. Some models combine aspects of both bipole and network scheme (Grossberg & Mingolla, 1985).

Current models do not implement all that is known about the geometry of relatability or certain other influences on contour interpolation, such as support ratio (but see Grossberg, Mingolla & Ross, 1997). Yet, the basic concepts of cooperative (e.g. bipole) units and network interactions seem likely to be the building blocks of future advances. A virtue of both approaches is their use the outputs of orientation-sensitive cells in early cortical visual areas. The existence of extensive recurrent connections (connections feeding back from later levels to earlier ones) in the relevant cortical areas has been argued to support hybrid models (Grossberg & Mingolla, 1985) although other models (Heitger et al., 1998) have stressed purely feed-forward processing as a virtue. (For further discussion of the relation of geometric and neural models in object formation, see Kellman et al., 2001).

A controversial argument about the process as described above is the *identity hypothesis*—the claim that the same interpolation process applies to several different looking contour interpolation phenomena. These include the phenomena in Fig. 1. Occluded objects, illusory objects, and other cases of interpolation differ in appearance and involve differing assignments of depth and boundary ownership relative to adjacent surfaces. Despite these differences, in appearance, research indicates that an identical interpolation process operates in all of these cases (Kellman, Yin, & Shipley, 1998; Ringach & Shapley, 1996; Shipley & Kellman, 1992a). Some models (Grossberg & Mingolla, 1985) and recent evidence (Guttman & Kellman, 2000; Kellman et al., 2001) suggest a *promiscuous* early interpolation process that makes more connections than are realized in the final scene representation. Later

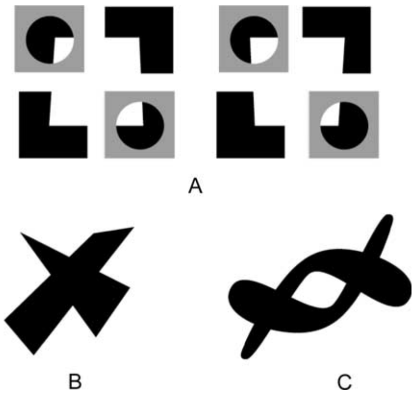

Fig. 4. (A) Illusory and occluded contours join. Display is a stereo pair; free-fuse by crossing the eyes. (B) and (C) Examples of homogeneous areas that split into two perceived objects (see text).

stages delete some interpolations and determine final depth arrangements of interpolated and other contours and surfaces.

Fig. 4 shows two phenomena that implicate a common interpolation process. Fig. 4A presents a stereo pair that may be free-fused by crossing the eyes to obtain a 3D display. (The effects to be described are also visible in each image alone, although less vividly due to all parts lying in a single depth plane.) The central white figure has 4 interpolated contours. Each one is an illusory contour for part of its length and an occluded contour along another part. That these *join* to form a single contour (or perhaps more to the point, that interpolation *occurs* between these two different kinds of endpoints) is significant and unexpected. It indicates that the interpolation process accepts either type of input, including combinations. This would be strictly prohibited in some models (Heitger et al., 1998), because of constraints on the inputs allowed to a given bipole operator. Fig. 4B (after Petter, 1956) shows a display having entirely homogeneous surface characteristics. This kind of display is an excellent demonstration and test of the basic segmentation and grouping concepts described above. Although one might expect a homogeneous region to be perceived as a single unified object, it splits into two perceived objects. The reason is the presence of the four tangent discontinuities in the concavities of the display, and the relatable edges leading into them. Once interpolation occurs, closed contours formed by real and interpolated edges define each of the two objects, the triangle and the quadrilateral.

This kind of display has some interesting properties relevant to the identity hypothesis. Where the two objects cross in the image, one appears in front, having illusory contours, and the other appears to pass behind, having occluded contours. With prolonged viewing of Fig. 4B, the depth order appears to switch, so that illusory contours become occluded ones and vice versa. There is an interesting regularity in such displays (Petter, 1956). Where interpolated boundaries cross, those crossing the smaller gap appear in front, as illusory contours, whereas the boundaries crossing the larger gap are seen behind, as occluded contours (Fig. 4C). The depth order ambiguity in Fig. 4B derives from the approximately equal lengths of interpolated edges. This claim about relative extent of interpolation, sometimes called Petter's Effect, has substantial experimental support and has been around for a long time. What has only recently been realized is that the effect provides a strong argument for an interpolation mechanism common to illusory and occluded contours. The reason is that in these displays, the determination of illusory vs. occluded appearance depends on a *comparison* of interpolated contours. Which appears in front depends on the relative lengths of crossing contours. If the illusory or occluded appearance of an interpolated contour depends on its length relative to a crossing interpolated contour, there must be, implicitly or explicitly, a length comparison process. This implies, in turn, that the locations and extents of contour interpolation are determined prior to their final appearance as illusory or occluded, at least in this case.

Several logical arguments of this kind as well as substantial experimental evidence combine to support the notion of a common underlying interpolation process, one that produces very precise and essentially identical representations of boundary locations in both illusory and occluded interpolation cases (Kellmam, Temesvary, Palmer, & Shipley, 2000; Kellman et al., 1998). Of course, different looking interpolation phenomena do involve differences in other aspects of scene representation, such as their depth relations with (projectively) adjacent surfaces.

The idea of a common perceptual process underlying illusory objects and occluded ones cuts against deep-seated assumptions in sensory and perceptual research. We will take up some of these issues in the last section.

3.2.4. Three-dimensional and spatiotemporal interpolation

Most interpolation research has addressed static, 2D displays. Some recent efforts have enlarged the domain to 3D and moving contours. It is clear that interpolation processes utilize 3D and motion-carried information and produce representations of 3D objects.

Fig. 5 shows 3D interpolation. Both rows show stereo pairs that should be free-fused by crossing the eyes.

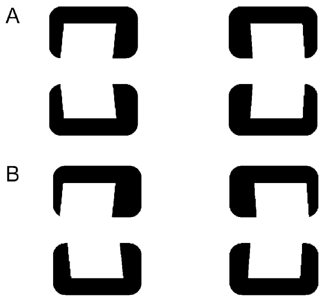

Fig. 5. 3D Interpolation displays (see text).

In Fig. 5A (and Fig. 4A above), illusory surfaces are visible. These utilize as inputs 3D orientations and positions of edges, and produce as outputs contours and surfaces traversing all three spatial dimensions. Fig. 5A and B shows one of the findings of our research (Kellman, Yin, Garrigan, Shipley, & Machado, 2003), namely that an expanded 3D principle of relatability appears to govern 3D interpolation. In Fig. 5A but not Fig. 5B, contours can be connected by a smooth, monotonic curve; an interpolated surface is seen in the former case, but not the latter. Also important for our perception of complete objects is dynamic information. Moving observers and moving objects produce changing patterns of occlusion. Our recent research (Palmer, Kellman, & Shipley, 1997) indicates that a notion of spatiotemporal relatability appears to govern which successively visible fragments get perceptually connected. It appears that briefly visible regions are encoded in a buffer, along with information about their velocity relative to the observer. The velocity information is used to extrapolate over time the positions of these stored fragments. As new contours and regions become visible, they are connected to the previously viewed regions following the geometry of spatial relatability, here applied to currently visible and previously stored regions in updated positions.

Results to date point to commonalities of interpolation processes in 2D, 3D, and dynamic object formation. Integrating these in more comprehensive accounts remains a research task of high priority.

3.3. Surface interpolation

Contour interpolation is not the only visual process that links visible areas; surface relations also play a role. In Fig. 6, note that the black circles in the display have no tangent discontinuities and therefore cannot participate in contour interpolation. Those circles that appear within interpolated edges of the surrounding black figure appear as holes. This is a manifestation of surface interpolation. In contrast, the black spot outside the figure appears as a spot on a surface. Likewise the white circles appear as holes through which the background is seen. Surface spreading behind occlusion links visible areas but provides little shape information; it is confined by real and interpolated contours. Visible regions are connected across gaps in the input based on the similarity of surface qualities (e.g. lightness, color, and texture). (For details and experimental work, see Yin, Kellman, & Shipley, 1997, 2000).

4. Separating processes in object perception

An important challenge in research on object perception is distinguishing processes that may be separate, functionally, computationally, and/or neurally. The challenge affects both the interpretation of empirical data and the construction of models. For models, separating processes improves understanding of what is going on and permits independent tests of assumptions and constructs. This becomes a high priority in the case of complicated models that incorporate large numbers of hypotheses, multiple processes, and interactions among them, perhaps including recurrent feedback and a combination of stimulus-driven processing and top-down influences (Grossberg, 1994). Although a model may simulate certain phenomena, it will also be imperfect in some respects. Tracking the sources of error and subjecting a particular theoretical construct to scientific test depend upon being able to isolate its effects. A companion problem arises empirically. Sometimes when outcomes of two experiments appear to conflict, the reason is that more than one process is at work. Different paradigms or stimuli may be differentially sensitive to the underlying processes.

In object perception, an important and instructive issue involves the contributions of relatively local, stimulus-driven processes and more global processes, involving familiarity, symmetry, or simplicity. The model we have described to this point has no influences from the latter. Interpolation based on contour relatability works equally well when the outcomes are familiar or unfamiliar, symmetric or asymmetric objects. Yet, global symmetry or familiarity have often been claimed to be important, perhaps beginning with the Gestalt notion of 'good form' (Wertheimer, 1923/1958). Kanizsa (1979) analyzed a number of cases in which local continuity conflicted with global symmetry, concluding that the latter is a weak or questionable contributor to object formation.

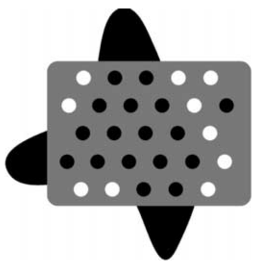

Fig. 6. Example of surface interpolation process that complements contour interpolation (see text).

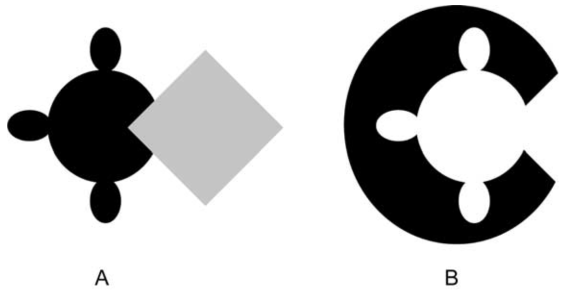

Fig. 7. (A) Occlusion display sometimes claimed to support global (symmetric) interpolation. (B) Illusory contour version of occlusion display in (A). No global completion is seen.

The controversy about local vs. global effects has continued, however. Evidence from priming studies has been consistent with both local and global effects (Sekuler, Palmer, & Flynn, 1994; van Lier, van der Helm, & Leeuwenberg, 1995). Fig. 7A shows an example. Global completion entails seeing a fourth articulated part behind the occluder, making the display radially symmetric.

In the boundary interpolation process as I have described it here, there is no way to create a fourth lobe behind the occluder. Compare Fig. 7B. In this illusory contour version of the display in Fig. 7A, there is no hint of a fourth lobe. Indeed, no illusory contour studies have suggested completions based on global symmetry in this sort of situation.

Yet one can still make sense of the claim of some observers to see the symmetric completion in Fig. 7A. Specifically, in an occluded display, observers are aware that part of an object is hidden from view. Apart from relatively low-level contour interpolation processes, cognitive processes may allow the observer to recognize that the visible areas are consistent with some familiar or symmetric figure. We have called this process recognition from partial information (RPI). RPI is conceptually distinguishable from processes that fill in contours. For example, a view of a helicopter's tail protruding from behind a building is sufficient to activate a representation of a helicopter, although its hidden boundaries may not be specified by available information. RPI would be consistent with the fact that global effects seem to be reported primarily in priming paradigms, as priming is known to be influenced by processes at various levels, including high level cognitive ones (e.g. in bilinguals, the English word 'bread' can prime the French word for butter).

Reasoning in this way, we predicted that boundary interpolation and RPI might be distinguishable experimentally. The former might lead to representations of precise boundary location in hidden regions, whereas the latter might not. To measure the precision of boundary location, we used a dot localization paradigm in which an occluded display was presented, followed by a brief probe dot in front of the occluder, after which the display was masked. Subjects were instructed to respond on each trial whether the probe dot fell inside or outside the occluded object's boundaries (i.e. whether the projection of the occluded object to the eye would or would not encompass the dot).

Using adaptive staircase procedures (stimulus values change over trials depending on the subject's responses), we were able to estimate the 0.707 probability of seeing the dot as outside the boundary and 0.707 probability of seeing the dot inside the boundary (= 0.293 probability of outside), respectively. *Imprecision* was measured as the *difference* between these estimates, and accuracy as their *mean*.

Staircases for several stimulus patterns were interleaved, and screen position varied randomly.

Kellman, Temesvary, Palmer and Shipley (2000, 2003) tested occluded displays similar to those used in priming research (Sekuler et al., 1994). Different groups of subjects were instructed that the occluded display should be completed globally or locally in order to induce in each case the best possible performance.

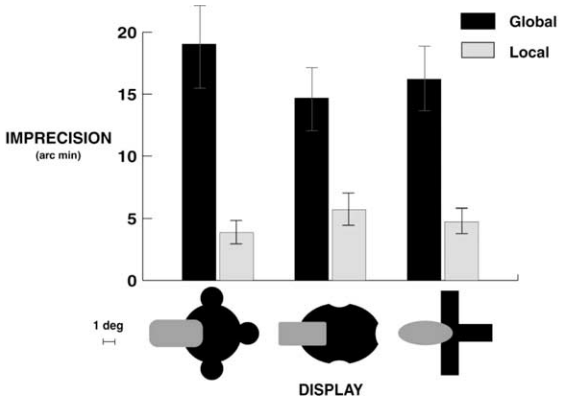

Fig. 8. Data from dot localization experiment. Imprecision in localization for the hidden boundary of each occluded figure is plotted, for global and local completion conditions (from Kellman et al., 2003).

Results showed that localization of boundaries in displays where completion is predicted by contour relatability was precise and accurate. Where completion is predicted to follow global symmetry, precision was far worse. Fig. 8 shows imprecision for several displays for both groups. Error in accuracy showed a similar pattern; subjects' judgments of the positions of boundaries predicted by global processes was very poor. This research suggests that global recognition processes may be experimentally separated from local boundary perception processes on the basis of the sorts of representations these create. Precise spatial positions of object boundaries are created in representations derived from latter process but not the former. These outcomes are consistent with the idea of separate perceptual completion and more cognitive RPI processes. Application of the dot localization method to illusory contour displays shows similar results: that real contours and illusory contours support good precision and accuracy of localization, whereas mere positional markers or displays with rounded tangent discontinuities do not (Guttman & Kellman, 2002).

5. Objects from interpolation: illusions, realities or representations?

What does it mean to say we *perceive* occluded regions of contours, surfaces, and objects? Do we really see them, or are they inferred? The issues are interesting, and they raise important questions about illusion, reality, and what it means to see.

The identity hypothesis, arguing for a common process in illusory and occluded contour completion, raises these issues directly, as these phenomena have traditionally been considered very different. Michotte (Michotte, Thines, &

Crabbe, 1964) labeled them as *modal* and *amodal* varieties of completion. Modal referred to the presence of sensory attributes, or modes, in the completed object, whereas amodal meant that the completed parts exist perceptually but without sensory attributes.

This distinction in phenomenology has most often led to a common sense conclusion. Illusory contours and surfaces are clearly *perceptual* phenomena, because we *see* them, i.e. they produce sensory experiences. Occluded contours and surfaces are out of sight, so in what sense can we be said to see them? They are *cognitive* phenomena; occluded parts are somehow known or inferred, not perceived.

Obviously, the account we sketched above clashes with this view of seeing vs. knowing. Where does the problem lie, in our unified account or in implicit conceptions of what it means to see?

The source of the problem, I believe, is confining the meaning of *seeing* to local sensory responses based on local physical data. Perceptual descriptions are functions of incoming information, but they need not be restricted to some narrow class of functions, such as a one-to-one correspondence between represented properties of objects and local sensations or stimuli.

This point has been made convincingly about once a generation in visual science, but it remains a source of abrasion between sensory and physiological approaches on one hand, and approaches to higher-order perception on the other. The Gestalt psychologists (Koffka, 1935; Wertheimer, 1958) presented a compelling case that perception is not a summation of local sensory responses. Gibson (1979) argued that sensations are actually quite incidental to perception. Rather than reviewing their arguments, perhaps the best way to sweep away the cobwebs here is to remind ourselves of the arguments for computation and representation in perception and information processing generally (Marr, 1982). On the computational/representational theory of mind, what is the principled difference between representing some contour as going behind another surface and representing some contour in front of another surface? Nothing. One is not more privileged or ‘real’ than the other. Characterizing one as a ‘sensory’ effect and the other a ‘cognitive’ effect reflects an epistemology that did not work out. Occluded contours are no more and no less ‘inferred’ than are illusory ones.

Given that there is no theoretical barrier separating represented contours and surfaces that are occluded or illusory, why do they *look* so different? What causes the robust difference in phenomenology? An answer is that this aspect of our phenomenology—the modal/amodal appearance—may simply code whether some surface or contour is nearest to the observer in some visual direction. This piece of information is important and, unlike the connectivity of objects, depends on the vantage point of the observer. Parts of an object that are nearest in a visual direction may be reached or grasped, whereas those equally real edges and surfaces behind some other object may not be reached without going around or through something.

The point is that all seeing is representational. So-called ‘real’ contours must be represented to be seen just as much as illusory or occluded ones. As we have seen, there is no basis for designating representations derived solely from local sensations as being the only ‘real’ ones. There are, however, multiple aspects of scenes that we need to represent. The connectivity of objects is one, best served by processes that are relatively indifferent to whether part of a contour goes behind or in front of something else. On the other hand, the reachability of a surface is a different aspect; it depends on the vantage point of the observer, and it is related to the phenomenology of modal/amodal. In the mind, there are not real contours, perceived contours, and inferred contours, only represented ones.

So, are interpolated parts of objects illusory or real? Our analysis brings us to a useful answer. For the representations needed for effective thought and action, encoding visible fragments alone would never do. Although faithful to local stimulation, such fragments are not the objects of the physical world. Interpolation processes pick up on basic regularities in the physical environment, probably as a result of evolutionary processes. As a result, they supply missing parts that lead to less faithful representations of the momentary image, but more accurate representations of the physical environment. In this sense they produce not illusions, but representations that bring us closer to reality.

Acknowledgements

Preparation of this paper was supported in part by a research grant from the National Eye Institute (R01 EY13518) to PJK. I thank Stephen Grossberg and an anonymous reviewer for comments on an earlier draft and Jennifer Vanyo for general assistance.

References

Banton, T., & Levi, D. M. (1992). The perceived strength of illusory contours. *Perception & Psychophysics*, *52*, 676–684.

Campbell, F. W., Cooper, G. F., & Enroth-Cugell, C. (1969). The spatial selectivity of the visual cells of the cat. *Journal of Physics*, *203*, 223–235.

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

Fischer, B., & Poggio, G. F. (1979). Depth sensitivity of binocular cortical neurons of behaving monkeys. *Proceedings of the Royal Society of London B*, *1157*, 409–414.

Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, *41*, 711–724.

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

Grossberg, S. (1994). 3D vision and figure-ground separation by visual cortex. *Perception & Psychophysics*, *55*(1), 48–120.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychoanalytic Review*, *92*, 173–211.

Grossberg, S., Mingolla, E., & Ross, W. D. (1997). Visual brain and visual perception: how does the cortex do perceptual grouping? *Trends in Neurosciences*, *20*(3), 106–111.

Guttman, S. E., & Kellman, P. J. (2000). Seeing between the lines: contour interpolation without perception of interpolated contours. *Investigative Ophthalmology and Visual Science*, *41*(4), S439.

Guttman, S. E., & Kellman, P. J. (2002). Do spatial factors influence the microgenesis of illusory contours? [Abstract]. *Journal of Vision*, *2*(7), 355a http://journalofvision.org/2/7/355/, DOI 10.1167/2.7.355.

Heitger, F., Rosenthaler, L., von der Heydt, R., Peterhans, E., & Kübler, O. (1992). Simulation of neural contour mechanisms: from simple to end-stopped cells. *Vision Research*, *32*, 963–981.

Heitger, F., von der Heydt, R., Peterhans, E., Rosenthaler, L., & Kübler, O. (1998). Simulation of neural contour mechanisms: representing anomalous contours. *Image and Vision Computing*, *16*, 407–421.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physics*, *195*, 215–243.

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

Kellman, P. J., & Arterberry, M. (1998). *The cradle of knowledge*. Cambridge, MA: MIT Press.

Kellman, P. J., Guttman, S., & Wickens, T. (2001). *Geometric and neural models of object perception. From fragments to objects: segmentation and grouping in vision (pp. 183–245)*, Amsterdam: Elsevier.

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

Kellman, P. J., Temesvary, A., Palmer, E. M., & Shipley, T. F. (2000). Separating local and global processes in object perception: evidence from an edge localization paradigm. *Investigative Ophthalmology and Visual Science*, *41*(4), S741.

Kellman, P.J., Temesvary, A., Palmer, E.M., & Shipley, T.F. (2003). Separating local and global processes in object perception, in preparation.

Kellman, P.J., Yin, C., Garrigan, P.B., Shipley, T.F., & Machado, L.J. (2003) Visual interpolation in three dimensions, in preparation.

Kellman, P. J., Yin, C., & Shipley, T. F. (1998). A common mechanism for illusory and occluded object completion. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 859–869.

Koffka, K. (1935). *Principles of Gestalt Psychology*. NY: Harcourt, Brace & World.

Marr, D. (1982). *Vision*. San Francisco: W.H. Freeman.

Michotte, A., Thines, G., & Crabbe, G. (1964). *Les complements amodaux des structures perceptives*. Louvain, Belgium: Publications Universitaires de Louvain.

Palmer, E. M., Kellman, P. J., & Shipley, T. F. (1997). Spatiotemporal relatability in dynamic object completion. *Investigative Ophthalmology and Visual Science*, *38*(4), S256.

Petter, G. (1956). Nuove ricerche sperimentali sulla totalizzazione percettiva. *Rivista di Psicologia*, *50*, 213–227.

Ringach, D., & Shapley, R. (1996). Spatial and temporal properties of illusory contours and amodal boundary completion. *Vision Research*, *36*, 3037–3050.

Sekuler, A. B., Palmer, S. E., & Flynn, C. (1994). Local and global processes in visual completion. *Psychological Science*, *5*, 260–267.

Shipley, T. F., & Kellman, P. J. (1992a). Perception of partly occluded objects and illusory figures: evidence for an identity hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 106–120.

Shipley, T. F., & Kellman, P. J. (1992b). Strength of visual interpolation depends on the ratio of physically-specified to total edge length. *Perception & Psychophysics*, *52*, 97–106.

van Lier, R. J., van der Helm, P. A., & Leeuwenberg, E. L. J. (1995). Competing global and local completions in visual occlusion. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 571–583.

Wertheimer, M. (1958). Principles of perceptual organization. In D. C. Beardslee, & M. Wertheimer (Eds.), *Readings in perception*. Princeton, NJ: Van Nostrand, (Original work published 1923).

Yin, C., Kellman, P. J., & Shipley, T. F. (1997). Surface completion complements boundary interpolation in the visual integration of partly occluded objects. *Perception*, *26*, 1459–1479.

Yin, C., Kellman, P. J., & Shipley, T. F. (2000). Surface integration influences depth discrimination. *Vision Research*, *40*, 1969–1978.