AMODAL COMPLETION: A CASE STUDY IN GROUPING

Allison B. Sekuler¹ and Richard F. Murray²
¹Department of Psychology, McMaster University
²Department of Psychology, University of Toronto

INTRODUCTION

Object perception is amazingly robust. We have little difficulty recognizing objects, despite the fact that retinal information about those objects varies dramatically from one moment to the next, and is often incomplete. One of the most common obstacles the visual system must overcome in order to recognize objects accurately is the partial occlusion of object contours and surfaces: Because most objects are opaque, they frequently occlude parts of themselves and parts of neighboring objects. Somehow, though, the visual systems seems to fill in missing information, so that we can interact appropriately with complete, meaningful objects.

In recent years, much research has focused on how the brain reconstructs this missing information, a process known as amodal completion. In some cases, researchers have been interested in the problem of amodal completion per se. However, we view amodal completion as a case study of the grouping and perceptual organization processes that are critical for object recognition more generally. First, as already mentioned, occlusion is ubiquitous, yet phenomenologically we seem to complete objects immediately and effortlessly. The completion of partly occluded objects is fundamental to object recognition, and because objects are so often occluded, the visual system has presumably developed efficient strategies to deal with the grouping of partly occluded contours. These same strategies may apply to other types of
grouping as well. Second, partly occluded objects are structurally ambiguous: There are an infinite number of possible completions for any partly occluded object. By understanding how observers resolve the structural ambiguity of partly occluded objects, we can determine the relative importance of various perceptual organization factors for object perception. Third, as will be discussed shortly, partly occluded objects have a long microgenetic time course relative to unoccluded objects. In other words, the visual system takes a relatively long time to arrive at a final visual representation for partly occluded objects. This elongated time course allows us to look more closely at the representational evolution of objects, and at the factors that influence that evolution.

During the past decade, there has been an explosion of research on visual completion, and this chapter is not intended to be a comprehensive review of all the research in the field. We focus primarily on the completion of two-dimensional static contours and surfaces, rather than on volumetric completion, completion of kinetic contours, or the relationships between amodal and modal completion. Other chapters in this volume speak more directly to some of those issues (e.g., Bruno; Kellman, Guttman & Wickens; Shipley & Cunningham; Shiffrar; van Lier). The goal of this chapter is to provide a selective overview of some empirical results concerning the visual completion of partly occluded objects, providing a summary of the current state of research and pointing out areas where more research is needed. Furthermore, the chapter is written with an eye toward providing strong constraints on the development of biologically plausible models of visual completion in particular, and of grouping more generally. The basic outline of the chapter is as follows. First, we give a brief historical overview of the empirical methods used in the study of visual completion, and we summarize the attempts researchers have made to establish the reality and importance of visual completion. We then describe some recent work that has begun to determine the physiological and psychological levels at which completion occurs, and we discuss the effects of spatiotemporal context, defined in a broad sense, on the time course, strength, and form of completion. Finally, we discuss the effects of amodal completion on processing efficiency and on observers' strategies in discriminating the shapes of unoccluded and occluded stimuli, and the implications of these results for developing biologically plausible, computational models of visual completion and grouping.
Gibson (1966, p. 203) noted that the everyday phenomena of occlusion are so familiar to us that we do not realize the need for an explanation. Indeed, the phenomenology of amodal completion is so compelling that many psychologists initially seemed to be satisfied with demonstrations based purely on appeals to subjective evidence (e.g., Kanizsa, 1979; Kanizsa & Gerbino, 1982). For example, experimenters might show observers a scene like Figure 1a and ask them to either draw or describe what completion they perceive, if any. A typical descriptive response is illustrated in Figure 1b ("A square behind a circle"), and a typical drawn response is illustrated in Figure 1c. Although the draw/describe approach dominated the study of amodal completion for quite some time, the information one can obtain from such studies is clearly limited by the nature of the task. First, the task is doubly subjective: it is subjective in that the observer must decide what response to make, based only on introspections that are accessible to no one else, and also in that the experimenter must interpret the response. For example, in the drawing illustrated in Figure 1c, one experimenter might interpret the lack of a sharp corner in the completed square as theoretically important, whereas another might interpret the same feature as a sign of laziness on the part of the subject. Second, in draw/describe tasks it is difficult to limit the time for which the observer processes the stimuli. This in turn makes it

![Illustration of the "describe" and "draw" paradigms. The observer views a stimulus as in (a), and either (b) describes or (c) draws the completion that he or she perceives.](image-url)
impossible to explore whether the internal representation of an occluded object changes over time – an issue that is critical for visual completion, as we discuss below.

These limitations of the draw/describe paradigm do not mean that phenomenology has no role in vision science. On the contrary, phenomenology serves a critical function in helping us frame important questions. Indeed, the modern study of completion began with the keen observational skills of Kanizsa and his colleagues. However, there are clear limitations to what phenomenology can reveal about the mechanisms underlying completion. Fortunately, over the past few decades, converging evidence from behavioral, physiological, and computational approaches has enabled us to begin to answer some important questions: In what objective sense does completion occur? What processes guide completion? How do those processes unfold over time? What neural systems control those processes?

A wide range of behavioral techniques and tasks have proved useful in recent years, including speeded matching tasks (Gerbino & Salmaso, 1985; Shore & Enns, 1997), adaptation (Brown & Weisstein, 1991), visual search (Davis & Driver, 1997; He & Nakayama, 1992;
Amodal completion

Rauschenberger & Yantis, 2001; Rensink & Enns, 1998), texture segregation (He & Nakayama, 1994c), primed matching (Bruno, Bertamini, & Domini, 1997; Guttman & Sekuler, submitted; Sekuler, 1994; Sekuler & Palmer, 1992; Sekuler, Palmer, & Flynn, 1994; van Lier & Wagemans, 1999), pattern discrimination (Gold, Murray, Bennett, & Sekuler, 2000; Murray, Sekuler, & Bennett, in press; Nakayama, Shimojo, & Silverman, 1989; Ringach & Shapley, 1996), dot localization (Kellman, Temesvary, Palmer, & Shipley, 2000; Takeichi, Nakazawa, Murakami, & Shimojo, 1995), object-based attentional cueing (Behrmann, Zemel, & Mozer, 1998; Moore, Yantis, & Vaughan, 1998; Pratt & Sekuler, in press), motion discrimination (He & Nakayama, 1994a; He & Nakayama, 1994b; Joseph & Nakayama, 1999; Sekuler & Sekuler, 1993; Sekuler & Sekuler, 1999; Yantis, 1995), and visual pursuit (Stone, Beutter, & Lorenceau, 2000). Variants of these tasks have also been adapted for the study of completion at the physiological level (Corballis, Fendrich, Shapley, & Gazzaniga, 1999; Giersch, Humphreys, Boucart, & Kovacs, 2000; Kersten, Shen, Ugurbil, & Schrater, 1999; Sugita, 1999).

Results from studies using these approaches support and extend results from phenomenological studies. For example, in the **primed matching paradigm** (Figure 2), observers view a priming stimulus and then judge whether a pair of test stimuli have the same shape or different shapes. The time taken to correctly identify "same" pairs depends on the representational similarity of the test shapes to the prime (Beller, 1971; Rosch, 1975a; Rosch, 1975b). For example, if observers are primed with a circle, they will be faster to respond "same" to a test pair of circles than to a test pair of 3/4-notched circles (Figure 2a). Conversely, if the prime is a 3/4-notched circle, observers will be faster to respond "same" to a pair of 3/4-notched circles than to a pair of circles (Figure 2b). This pattern of results is referred to as **basic shape priming**. With regard to occlusion, one interesting question is what happens when the priming stimulus is partly occluded, as in Figure 2c. Here, the prime we are interested in is the lower of the two shapes. Phenomenologically, observers describe the prime as a circle partly hidden behind a square (a complete representation). However, the pattern of information that reaches the eye is also consistent with the interpretation that a 3/4-notched circle is nudged up against the square (the mosaic representation). Sekuler and Palmer (1992) found that under some conditions, the pattern of priming for occluded objects is similar to that of complete objects, and quite different from that of mosaic objects. This result may be taken as objective, albeit indirect
evidence that the visual system treats occluded objects as if they were complete. We call this approach objective, in contrast to the draw/describe paradigm, because in the primed matching paradigm we do not ask observers about their subjective, introspective impressions of an occluded object: The inference that the visual system treats a partly occluded object as complete is based entirely on the effects of priming on observers’ "same" and "different" responses to pairs of test shapes. This paradigm also has the advantage that it can be used to explore the developmental time course, or microgenesis, of visual completion. By presenting the prime for varying amounts of time while holding the interstimulus interval fixed, we can take snapshots of the stimulus representation at different times in its development. Using this technique, Sekuler and Palmer showed that completion requires a measurable amount of time. With long prime durations (100-200 ms), partly occluded objects primed observers’ “same” and “different” responses like complete objects; but with short prime durations (50 ms) partly occluded objects primed observers’ responses like mosaic objects, or like an intermediate representation. As we discuss below, the primed-matching paradigm has proven extremely versatile, and has been used to address a wide variety of issues concerning amodal completion (e.g., Bruno et al., 1997; Guttman & Sekuler, submitted; Sekuler, 1994; Sekuler et al., 1994; van Lier & Wagemans, 1999). However, the primed-matching paradigm does have its limitations. First, basic shape priming is difficult to detect for short stimulus durations, so the technique is not sensitive to very rapid time courses. Second, basic shape priming can be difficult to detect reliably when response times are short, so individual observers cannot be tested extensively because practice speeds response times. Finally, even in conditions where it is appropriate to use the primed-matching paradigm, it is important to show that the results are not simply an artifact of one particular experimental method.

Fortunately, other converging methods, such as shape discrimination tasks, do exist. Like the primed-matching paradigm, shape discrimination tasks provide researchers with an objective method of determining how an occluded object is processed by the visual system, and also allow us to explore the microgenesis of completion (e.g., Murray et al., in press; Ringach & Shapley, 1996). For example, Figure 3 illustrates the stimuli and task used by Murray et al. (in press). In one condition, the shapes to be judged were complete (left panel); in another, the contours defining the shapes were fragmented (right panel); and in a third condition, the contours were
also fragmented, but occluders were present in the gaps between fragments (center panel). Phenomenologically, observers reported the occluded shape to be much more similar to the complete shape than to the fragmented shape. Murray et al. obtained a more objective measure of the similarities among these shapes by comparing performance on a shape discrimination task. Observers were asked to make a simple shape judgement: Is the rectangle longer vertically or horizontally? One would expect observers to be quite good at this task for complete stimuli, because observers are highly sensitive to small deviations from perfect symmetry in quadrilateral figures (Regan & Hamstra, 1992). One would also expect performance to be markedly impaired in the fragmented condition because much of the stimulus has been deleted, and because observers were required to distribute their attention over several perceptual units rather than focusing on a single object (Baylis & Driver, 1993; Behrmann et al., 1998; Duncan, 1984). The critical question in this task is how observers perform in the occluded condition, compared to the other two conditions. If the visual system treats the occluded object as though it is complete, then performance should be quite good. However, if the visual system treats the occluded object as a group of unrelated line segments, as in the fragmented condition, then performance should be poor. Like Sekuler and Palmer (1992), Murray et al. found that the results depended on the stimulus duration. For short durations (15-30 ms), performance in the occluded condition was as poor as performance in the fragmented condition. For longer durations (75 ms and longer), performance in the occluded condition was almost as good as performance in the complete condition. Hence their results supported the idea that although the visual system does in fact treat partly occluded objects as though they are complete, completion requires a measurable amount of time.
Figure 3. Stimuli from Murray et al. (in press). In each condition, the observer's task is to determine whether the white target pattern is longer vertically or horizontally (in all the stimuli shown, the correct answer is "horizontal"). The target pattern is either a complete rectangle, an occluded rectangle, or a fragmented pattern that contains the same visible line segments as the occluded rectangle, but lacks the T-junctions that give the impression that the line segments belong to a single partly occluded rectangle.

**CONTEXT EFFECTS**

Although studies have generally been consistent in concluding that completion requires time, the estimate of precisely how much time is required has varied considerably from one study to the next. Sekuler and Palmer (1992) estimated that 100-200 ms was required for completion; Ringach and Shapley's (1996) estimates ranged from 120-170 ms; Murray et al.'s (in press) estimate was approximately 75 ms; Guttman and Sekuler's (submitted) estimates ranged from less than 75 ms to over 400 ms; and Bruno et al. (1997) reported finding no measurable minimum time for completion. Clearly, completion does not require a fixed amount of time, but varies considerably depending on the context of completion.

We define context in the broadest sense, as the spatial, temporal, and attentional circumstances surrounding the stimulus to be completed. As such, context includes the size and shape of a stimulus, the presence or absence of additional cues for depth and grouping beyond occlusion cues, and the recent history of an object (e.g., as described by an object file Kahneman, Treisman, & Gibbs, 1992). We are only beginning to understand the full extent to which these contextual factors affect visual completion, but such an understanding is crucial for the
development of accurate, biologically plausible computational models of grouping. Here we describe some results that constrain the development of such models.

**Context affects completion time**

As described above, it is well established that amodal completion takes time, but estimates of the time required for completion are variable. Two factors that play a role in determining time to completion are (1) the amount of occlusion, and (2) the presence of depth cues other than occlusion cues.

**Amount of occlusion.** As the amount of occlusion increases, completion could be affected in two ways, described by the *temporal limit hypothesis* and the *spatial limit hypothesis* (Guttman & Sekuler, submitted; Shore & Enns, 1997). The temporal limit hypothesis states that visual completion is constrained by temporal factors. Specifically, as the amount of occlusion increases (as in Figure 4, top panel), so too should the time required for completion. The temporal limit hypothesis is consistent with models of completion based on the propagation of contours (e.g., Grossberg & Mingolla, 1985; see Mingolla, this volume, for details).
Alternatively, visual completion may be constrained by spatial factors, regardless of processing time. The spatial limit hypothesis states that as the amount of occlusion increases beyond some limit, completion no longer occurs. Intuitively, this hypothesis makes sense: In the limit of total occlusion, there is no visible object for the visual system to complete. The theoretical support for the spatial limit hypothesis comes mainly from the influential relatability theory of Kellman and Shipley (1991; 1992). This theory, components of which have received much empirical support (Field, Hayes, & Hess, 1993; Kellman, Yin, & Shipley, 1998; Shipley & Kellman, 1990; Shipley & Kellman, 1992), suggests that completion occurs only when the visible segments of occluded contours are relatable, meaning that the tangents to the contours at points of occlusion meet at an angle of no more than 90 degrees, and can be connected by a curve containing no inflection points or discontinuities. This theory can be seen as a modern instantiation of the Gestalt notion of good continuation: Local regions of occlusion are identified based on the presence of T-junctions (Helmholtz, 1910/1962; Ratoosh, 1949), and the contours bounding the occluded surfaces are connected by the simplest continuous contour. Consider, for example, a circle partly occluded by a square (Figure 4, top panel). When 25% or less of the circle's contour is occluded, the occluded contours are relatable. When more than 25% of the circle's contour is occluded, the occluded contours are not relatable, and completion should not occur.

Several studies have investigated how the amount of occlusion affects completion (Guttman & Sekuler, submitted; Rauschenberger & Yantis, 2001; Shore & Enns, 1997). The results of these experiments provide converging evidence that supports the temporal limit hypothesis. Shore and Enns showed observers scenes containing complete shapes, occluded shapes, and mosaic shapes. For each image, observers made a speeded judgment about whether a target shape was a (possibly notched) circle or square. Observers were fast at correctly classifying complete shapes, and were slow at correctly classifying mosaic shapes. However, the time taken to correctly classify occluded shapes varied as a function of the amount of occlusion. For shapes occluded by less than 25%, classification times were similar to those for complete shapes, whereas for shapes occluded by more than 25%, classification times were much slower, and statistically indistinguishable from those for mosaic shapes. Although this result could be seen as supporting the spatial limit hypothesis (because no evidence of completion was found for
occlusion over 25%), it is also consistent with the temporal limit hypothesis (because the visual system may simply require more time than was allowed for completion of highly occluded objects).

Using a visual search paradigm, Rauschenberger and Yantis (2001) also found results consistent with changes in completion as a function of the amount of occlusion. Following previous researchers (Davis & Driver, 1997; He & Nakayama, 1992; Rensink & Enns, 1998), Rauschenberger and Yantis assumed that if the visual system represents an occluded circle as a mosaic (i.e., as a notched circle), then it should "pop out" of a set of distractors comprising complete circles, and search times should not vary much as a function of the distractor set size (so-called “parallel search”, e.g., Palmer, 1995; Treisman, 1988; but see Palmer (1995)). Conversely, if the occluded circle is represented as complete, then it should be quite difficult to detect among a field of circle distractors, and significantly larger set-size effects should be found (so-called "serial search", Treisman, 1988). Rauschenberger and Yantis found that whether occluded circles popped out of circle backgrounds depended on both the amount of occlusion and the time for which the display was viewed before masking. When circles were occluded by less than 10%, search was "serial" even when stimuli were displayed for only 100 ms. When circles were occluded by 25%, search was "parallel" when stimuli were presented for 100 ms, but search was "serial" when stimuli were presented for 250 ms. Finally, when circles were occluded by 37%, search was parallel for both 100 ms and 250 ms presentations. These results are consistent with the idea that time to completion increases as a function of the amount of occlusion, at least up to 25%, which is the limit predicted for these stimuli by Kellman and Shipley's (1991; 1992) relatability hypothesis. Beyond this amount, Rauschenberger and Yantis found no clear evidence of completion, although the longest duration they tested (250 ms) may simply not have been long enough to reveal the effects of completion for a highly occluded object.

More recently, Guttman and Sekuler (submitted) used the primed-matching paradigm to test the hypothesis that even highly occluded objects can be completed, given enough time. In their study, the pattern of priming for occluded objects varied as a function of both the amount of occlusion in the prime and the amount of processing time for the prime. Figure 4 (bottom panel) summarizes the results from two of their conditions (20% contour occlusion and 32.5% contour...
occlusion), plotting the value of a completion index $c$ as a function of prime-to-test SOA. With their completion index, a value of $c=1$ indicates that priming of the occluded object was identical to that of complete objects, and a value of $c=0$ indicates that priming of the occluded object was identical to that of mosaic objects. Values greater than 0.8 on this scale are taken to indicate effective (though not perfect) completion. With 20% occlusion, $c$ reached this critical value for even the shortest SOA tested (75 ms). With 32.5% occlusion, $c$ once again reached the critical value, but only after a much longer time. This is a crucial result, because it shows for the first time that highly occluded object can be completed when given enough time, even when the condition of relatability is not met.

Taken as a group, these studies provide strong support for the temporal limit hypothesis. Thus, although spatial limits may exist as well, those limits are not defined by the condition of relatability – a fact that must be taken into account in the further development of computational models of completion and grouping. Guttman and Sekuler suggested that a modified notion of relatability may prove useful for understanding the role of spatial limits in visual completion. In their scheme, the curvature of contours leading into a discontinuity guides completion (Takeichi et al., 1995) in addition to the tangents at the points of occlusion (Kellman & Shipley, 1991; Kellman & Shipley, 1992). Guttman and Sekuler further suggested that the influence of relatability on visual completion may depend on other factors that affect how informative an object’s visible contours are about its occluded contours, such as the length of the contours leading to points of occlusion, and the consistency of their curvature. Thus, although the original notion of relatability cannot completely explain the results of the experiments described in this section, a modified notion of relatability in combination with other factors may well be an important component of future models.

Presence of additional depth cues. The amount of occlusion is not the only factor that can affect the time to completion of partly occluded objects. Recent behavioral and physiological work suggests that time to completion decreases with the addition of depth cues that are consistent with the depth ordering indicated by occlusion cues (Bruno et al., 1997; Sugita, 1999).

Bruno et al. (1997) correctly noted that in Sekuler and Palmer's (1992) primed-matching study, the occluder was presented in the same stereoscopic depth plane as the occluded figure. Thus occlusion and the resulting relative depth assignment of objects was specified only by
discontinuities (T-junctions) at points of occlusion. Although stereopsis plays little role in many viewing situations, it is certainly an important cue to depth for nearby objects. Thus it is important to understand the role that stereoscopic depth cues play in amodal completion. In an elegant experiment, Bruno et al. again used the primed-matching paradigm, but examined the effect of consistent stereoscopic depth cues. When no consistent stereoscopic depth cues were presented, Bruno et al. replicated Sekuler and Palmer's results: Some measurable time was required before an occluded prime yielded a pattern of priming similar to that of a complete prime. However, when Bruno et al. added stereoscopic depth cues that were consistent with the depth ordering indicated by occlusion, occluded primes yielded a pattern of priming similar to that of a complete prime, even at the shortest SOA tested (50 ms). Although one could interpret this result as indicating that completion required no time in the presence of consistent stereoscopic depth cues, a more plausible interpretation is that the process is simply speeded beyond the point where it can be detected using the primed matching paradigm (as noted earlier, the primed-matching paradigm is not ideal for revealing very rapid completion processes).

Recent physiological evidence points to a mechanism that could explain the reduced time to completion in the presence of consistent stereoscopic depth cues. Sugita (1999) found that some orientation selective neurons at the earliest level of processing in the visual cortex (V1) responded to partly occluded contours as though they were complete. This result suggests that neurons as early as V1 have the computational power to complete partly occluded contours. However, these neural responses are observed only when the occluder is presented stereoscopically in front of the occluded contour. If the occluder is presented in the same stereoscopic depth plane as the occluded contour, these same neurons cease to respond to the occluded contours. With respect to Bruno et al.'s (1997) results, this suggests that fast completion times in the presence of consistent stereoscopic depth information may have been mediated by early levels of visual processing in the cortex, perhaps as early as V1. On the other hand, slower completion times in the absence of consistent stereoscopic depth cues may have been mediated by later regions of visual cortex. It is difficult to evaluate this hypothesis, because relatively little physiological research has been done on amodal completion, compared to possibly related phenomena such as illusory contours (e.g., Grosof, Shapley, & Hawken, 1993; Sheth, Sharma, Rao, & Sur, 1996; von der Heydt, Peterhans, & Baumgartner, 1984). However,
at least one study suggests that shape selective neurons in the inferotemporal cortex show the same shape selectivity in the presence of occlusion (without stereoscopic depth cues) as when the shapes are fully visible, although the relevant neurons’ overall rate of responding decreased as a function of increasing occlusion (Kovacs, Vogels, & Orban, 1995). Interestingly, there also appeared to be an inverse correlation between the rate of responding and the time to peak firing (Vogels, personal communication). This result is qualitatively consistent with the effect of the amount of occlusion on time to completion, as discussed in the last section, although additional studies are clearly required to quantify this relationship more precisely.

Context affects completion strength. Just as the speed of amodal completion can vary from one context to another, so too can the asymptotic strength of completion vary. As described previously, Murray et al. (in press) found that at long stimulus durations, observers were approximately as good at judging the aspect ratio of a partly occluded rectangle (Figure 3, middle panel) as they were at judging the aspect ratio of a complete rectangle (Figure 3, left panel). Furthermore, at long stimulus durations observers were much better at judging the aspect ratio of an occluded rectangle than they were at judging the aspect ratio of fragmented rectangle (Figure 3, right panel). Murray et al. formulated a measure of amodal completion, similar to that described earlier, in which the strength of completion was $c=1$ (full completion) if observers’ aspect ratio discrimination thresholds were the same in the occluded and complete conditions, and the strength of completion was $c=0$ (no completion) if observers’ thresholds were the same in the occluded and fragmented conditions. They used this measure to investigate the time course of amodal completion, and in particular to determine whether motion cues affected the speed or strength of completion.

Murray et al. found that the speed of completion did not depend on whether the rectangle was in motion, even though under many circumstances motion is a strong cue for completion and grouping (Kellman & Shipley, 1991). Regardless of whether the rectangle was static or moving, Murray et al. found that completion occurred quickly, taking approximately 75 ms. However, the asymptotic strength of completion did depend on motion cues. If the rectangle was in motion, the strength of amodal completion rose rapidly and asymptoted at approximately 1.0, meaning that at long durations, thresholds were the same in the occluded and complete conditions. If the rectangle was stationary, the measure of completion rose rapidly and
asymptoted around 0.8 (similar to the values of completion strength found for static objects by Guttman & Sekuler, submitted). In other words, even for relatively long stimulus durations (up to 220 ms), thresholds were always slightly lower in the complete condition than in the occluded condition. Hence, Murray et al.’s results demonstrate that context can affect the asymptotic strength of completion, and that this effect can be dissociated from effects on the speed of completion.

**Context affects completion form**

Another way that context affects completion is by influencing how shapes are perceived to continue behind occluding surfaces. These effects have at least two sources: (1) the spatial structure of a scene can affect the form of completion, and (2) past perceptual events can also influence the form of completion.

**Spatial structure.** The most striking example of how different spatial structures can lead to different visual completions is illustrated by the debate over the relative roles of *local* versus *global* processes in completion. According to local models, completion is based only on attributes of contours and surfaces at or near points of occlusion. The most common instantiation of this model leads to a "good continuation" solution, in which contours are completed by the simplest possible continuous curve (e.g., Kellman & Shipley, 1991; Kellman & Shipley, 1992; see also the chapters in this volume by Kellman, Guttman, & Wickens, and by Shipley & Cunningham). One advantage of this type of model is that it is computationally tractable, and the hypothesized interactions among input units along a contour are consistent with the physiology of early visual cortical processing. Not surprisingly, such models have gained considerable momentum over the past decade.

According to global models, completion is based on attributes of surfaces and contours of the entire object to be completed (and in some cases on attributes of nearby objects as well). One example of a global model, although it is not a process model, is Structural Information Theory (SIT; Buffart & Leeuwenberg, 1981; Buffart, Leeuwenbert, & Restle, 1981; see van Lier, this volume, for an overview). SIT predicts a Prägnanz solution to the problem of visual completion: The globally simplest, most regular completion is favored over others. Although early phenomenological approaches supported global models (e.g., Dinnerstein & Wertheimer,
From Fragments to Objects

1957), it is unclear how global completions might be computed from known neuronal responses, and consequently less research has focused on global models.

For most of the stimuli described so far, the local and global completion solutions are identical (e.g., a circle completed behind a square). However, in more realistic stimuli, ambiguities are likely to occur, and local and global processes may lead to qualitatively different completions (e.g., Davi, Pinna, & Sambin, 1992). Recently, some of the objective experimental paradigms described earlier have been used with such ambiguous stimuli. The results of these studies suggest that, despite our difficulty in modeling global completion processes, global as well as local processes play a role in completion (Sekuler, 1994; van Lier & Wagemans, 1999; van Lier, Leeuwenberg, & van der Helm, 1995). Thus, these results support so-called "hybrid" models that suggest that completion is influenced by both the local and the global spatial structure of a stimulus (e.g., van Lier, van der Helm, & Leeuwenberg, 1994).

Sekuler and colleagues (1994) used the primed-matching paradigm to explore the conditions under which local or global processes dominate completion. In their stimuli (Figure 5), local models predicted one completion, and global models predicted another. Their results were consistent with previous reports based on subjective methods (e.g., Boselie & Leeuwenberg, 1986; Buffart & Leeuwenberg, 1981; Buffart et al., 1983; Dinnerstein &
Wertheimer, 1957), which found that global processes can in fact play a significant role in completion. In a follow-up study, Sekuler (1994) identified some of the factors that determine the influence of local and global processes, and proposed a qualitative model of completion in which costs and benefits lead to a weighted solution based on both local and global processes. In this weighted hybrid model, the costs of global solutions are relatively high compared to those of local solutions because local solutions are easily computed even at the earliest levels of the visual cortex, whereas global solutions are not. Consequently, the benefits must be significant for global processes to play a major role. For a specific example, consider the shapes in Figure 5. The local solution to the occluded shape in Figure 5a yields one-fold bilateral symmetry (about the vertical axis), whereas the global solution yields two-fold bilateral symmetry (about the vertical and horizontal axes). In this case, the additional costs of the global solution add only horizontal symmetry, which has been shown to be a significantly weaker form of symmetry than vertical symmetry (Corballis & Roldan, 1975; Goldmeier, 1972; Mach & Szasz, 1886/1959; Palmer & Hemenway, 1978; Pashler, 1990; Sekuler, 1994; Wagemans, Van Gool, & d'Ydewalle, 1991; see Wenderoth, 1996, for a recent review). Thus, the benefits here do not outweigh the costs, and results from the primed-matching paradigm suggest that local processes do in fact dominate completion for this stimulus (Sekuler, 1994). A much greater contrast is seen between the local and global solutions to the occluded shape in Figure 5b. Here, the local solution yields one-fold bilateral symmetry (about the vertical axis), but the global solution yields four-fold bilateral symmetry (about the vertical, horizontal, and two diagonal axes). Thus, the benefits here are larger than in the previous example, and results from the primed-matching paradigm suggest that the stimulus global processes do dominate completion of this stimulus (Sekuler, 1994). In general, a large weight will be assigned to the global solution when the global completion is much more regular than the local completion. For example, with completely irregular shapes we only ever expect to find evidence for local completion. It is important to note that because the high cost of a global solution is factored into this qualitative model, the global solution is not favored when it requires a very complex completion. For example, we do not expect the visual system to complete the intricate features of a face behind an occluder.

The notion that amodal completions are based on weighted combinations of different processes (the weighted hybrid model) has two important implications. First, it suggests that
when conflicts occur between local and global completion solutions, the internal representation of the completed object does not map directly onto one or the other solution, but is somewhere in between. Support for this idea can be found in the results of Sekuler (1994; see also van Lier et al., 1995). Sekuler found that although global processes dominated completion (i.e., occluded objects led to better priming for globally regular objects than for locally completed objects), the amount of priming could differ markedly between occluded primes and complete global primes. This result is consistent with the notion that the internal representation of the occluded object lies between the local and global solutions, but is biased toward the global solution.

A second implication of the weighted hybrid model is that the fidelity of internal representations should vary depending on context. For example, when local and global processes lead to the same solution, one would expect relatively stable representations. When local and global processes lead to conflicting solutions, one would expect less stable representations. In addition, one might expect greater stability for representations that have heavier weightings for local solutions, simply because of the low cost of local solutions compared to global solutions. Recent work using the dot-localization paradigm has provided support for this idea (Kellman et al., 2000). Kellman and colleagues presented observers with partly occluded shapes similar to those in Figure 5b. As described earlier, these shapes have different local and global solutions. In one condition, observers were asked to adopt a ‘local’ mindset, and to imagine what the object would look like if completed according to a local process. In another condition, observers were asked to adopt a ‘global’ mindset, and to imagine what the object would look like if completed according to global processes. In both conditions, after the observer viewed the occluded object for some time, a dot appeared superimposed on the occluder. Observers were asked to indicate whether the dot appeared inside or outside the imagined boundary of the completed object, and an error was computed relative to where the boundary actually would have been for either the local or global solution. Kellman et al. found that observers had much higher errors in the dot-localization task when they adopted the global mindset than when they adopted the local mindset. Based on these results, Kellman et al. suggested that the representation of the global solution is spatially less well-defined than that of the local solution. This interpretation is entirely consistent with the idea that representational fidelity varies with context, and in this case the context is defined by the spatial structure of the
Amodal completion

stimulus. The dot-localization technique may prove to be quite useful in determining finer scale properties of amodal (and other) completions than can be resolved using paradigms such as primed-matching. For example, we suggested above that even highly occluded, non-relatable contours can be completed when the visual system is given enough time. However, it may be that the fidelity of the completed representation, as assessed by the dot-localization paradigm, varies as a function of the amount of occlusion.

Past perceptual events. Even if the visual system uses a weighted hybrid method to complete occluded objects, it is likely that the relative weights assigned to the local and global processes of amodal completion vary with context. For example, an observer's past experience may be a critical factor for determining the relative weights of local and global processes in completion. This experience might include long term familiarity effects as well as short term priming effects. Consider a task in which observers perform some initial practice trials in which they judge whether two patterns have the same shape or different shapes. In different conditions, 80%, 50%, or 20% of the practice trials might contain highly symmetric shapes that suggest a global completion, with the remainder of the shapes less symmetric and suggesting a local completion. If following these practice trials the observer views occluded stimuli that could be completed either globally or locally, one might expect that observers who viewed a greater proportion of globally completed shapes in the practice session would assign greater weight to global processes. In a preliminary experiment using the primed-matching paradigm and only 100 practice trials, we did indeed find a trend in this direction. Additional work clearly is needed to determine the amount of practice that would produce a significant effect, and whether the effect of past experience depends on shape tokens or types (e.g., Kahneman et al., 1992).

Although additional work is required to determine the effect of previous experience on the shape of a completed contour, it is already clear that experience can influence the extent to which completion occurs. The best evidence of this comes from an experiment conducted by Joseph and Nakayama (1999). Joseph and Nakayama presented observers with an ambiguous apparent motion stimulus, in which motion could be seen in either the horizontal or vertical directions (Figure 6a). The motion is ambiguous in this stimulus because the distance between horizontal elements is the same as the distance between vertical elements. Previous work by Shimojo and Nakayama (1990) showed that when an occluder is placed in this scene (Figure 6b),
the stimulus becomes perceptually less ambiguous, and observers show a bias to perceive motion behind the occluder (i.e., vertically in Figure 6b). Shimojo and Nakayama suggested that this effect resulted from the perceptual continuation of an occluded object behind the occluder. This continuation decreases the effective distance between vertical elements, biasing the apparent motion in that direction. In their extension of this work, Joseph and Nakayama asked whether the amount of bias in this apparent motion task depends on how an observer has previously perceived the elements in the scene. In other words, does the recent context in which an observer has seen an object influence the completion of that object? In Joseph and Nakayama’s study, observers saw previews of the moving elements, whose size was either the same as elements abutting the occluding surface (short elements; Figure 6c) or whose size was consistent with elements continuing behind the occluding surface (long elements; Figure 6d). Later, when judging the perceived direction of the ambiguous apparent motion display, observers showed a greater bias to see motion behind the occluder if they had seen the longer preview elements. Previews of long elements increased the amount of continuation behind the occluder, and these
preview effects lasted for at least one second in all observers tested, and even longer in some subjects (at least six seconds in one case). These results show that the recent history of an object can influence the way in which the visual system interprets and processes occlusion cues.

It is important to note, however, that previous experience with an object does not always dominate stimulus-derived information. Using an object-based attention task adapted from Egly, Driver, and Rafal (1994), and Moore, Yantis, and Vaughan (1998), Pratt and Sekuler (in press) showed that a preview of a truncated object does not always prevent completion in the presence of strong grouping cues, even when the occluded stimulus is shown for a relatively short duration. This result shows that a full model of completion (and of grouping more generally) must incorporate interactions among internal factors (such as the effect of past experience) and stimulus-derived factors (such as stimulus structure).

**HOW DOES AMODAL COMPLETION AFFECT PERCEPTUAL PROCESSING?**

**What changes with completion: efficiency or internal noise?**

As we have just discussed, context can affect amodal completion in a variety of ways, and amodal completion can affect performance in a wide range of perceptual tasks. However, the studies described so far do not speak to the issue of how amodal completion affects performance. Performance in any perceptual task is constrained by two factors: (1) the efficiency with which an observer uses relevant information from a stimulus (i.e., calculation efficiency), and (2) random variations in the observer, such as random variations in sensory encoding (i.e., internal noise). In this context, the question becomes: Does amodal completion affect performance in perceptual tasks by changing calculation efficiency, by changing internal noise, or both? Noise masking experiments allow us to measure the influence of each of these factors on an observer’s performance in a perceptual task, and hence to characterize an observer’s performance in more detail than can be achieved by gross measures of performance, such as percent correct or discrimination thresholds (Burgess & Colborne, 1988; Pelli & Farell, 1999).

In a noise masking experiment, one measures discrimination thresholds in a task, in several levels of external white Gaussian noise. The function that plots discrimination thresholds
versus external noise power is called a *noise masking function*. Empirically, noise masking functions are always found to be approximately linear. A well-established body of results, largely derived from signal detection theory, shows that the x-intercept of the noise masking function can be regarded as a measure of the noise in the observer’s early encoding of the stimulus, and the slope gives a measure of how efficiently the observer uses the stimulus to perform the task (Burgess & Colborne, 1988; Pelli & Farell, 1999).

Recently, Murray and colleagues (2000) applied noise masking methods to a task developed by Ringach and Shapley (1996). In Ringach and Shapley's study, observers discriminated between ‘fat’ and ‘thin’ Kanizsa squares (left and right columns of Figure 7). Fat and thin stimuli are distinguished from one another by a small difference in the orientation of the notched circle inducers at the corners of the square. In one condition, observers discriminated between fat and thin stimuli that had luminance-defined contours connecting the inducers (real condition, Figure 7). In a second condition, observers discriminated between fat and thin amodally completed Kanizsa squares that looked like squares seen through four holes in an occluding surface (occluded condition, Figure 7). In a third condition, all the inducers faced in the same direction (fragmented condition, Figure 7). Locally, the occluded and fragmented stimuli were very similar, and were distinguished only by the orientations of the inducers. Ringach and Shapley showed that, at long stimulus durations, discrimination performance was similar in the real and occluded conditions, and significantly worse in the fragmented condition. It is understandable that performance in the real condition should be better than in the fragmented condition, because the stimulus in the real condition contains more contour that the observer can use to perform the task. But why are performance levels so different in the occluded and fragmented conditions? Here, the same amount of physical information is available in both tasks (eight luminance-defined inducer edges), so naively we might expect that performance would be roughly the same.

Murray *et al.*'s noise masking experiments answered this question with respect to the two fundamental limits to performance: calculation efficiency and internal noise. Murray *et al.* found that observers showed no consistent differences between the occluded and fragmented tasks in the amount of internal noise that limited their performance. However, observers did show large differences between the two tasks in terms of the efficiency of their decision strategy.
(calculation efficiency): Observers were much more efficient in the occluded condition than in the fragmented condition, and in fact they were at least as efficient in the occluded condition as in the real condition. Hence Murray et al.’s results answer the question of how amodal completion brings about such large changes in observers’ performance: Completion changes the efficiency of the computation that observers use to make their perceptual decisions, but does not change the levels of internal noise. This result strongly constrains the classes of models that can explain the role of amodal completion in visual processing, and the role of grouping more generally.

Mapping behavioral receptive fields for amodal completion

Murray et al.’s (2000) results show that amodal completion can affect the efficiency of observers’ strategies in perceptual tasks. However, their results say nothing about how observers’ strategies differed across the three tasks they examined. What is it that makes an observer more efficient in the occluded condition than in the fragmented condition? Ringach and Shapley (1996) suggested that the answer may involve the parts of the stimulus observers use to perform the task. For example, observers may have used all four sides of the amodally completed square to perform the task in the occluded condition, but only a single inducer in the

Figure 7. Fat and thin Kanizsa squares (left and right columns), and behavioral receptive fields (center column) for real, occluded, and fragmented stimuli. See text for details.
fragmented condition. This hypothesis can be tested using another noise masking technique: *response classification*.

In a response classification experiment, the observer performs a discrimination task on stimuli embedded in Gaussian white noise, and over many trials one measures the correlation between the noise contrast at each stimulus pixel and the observer’s responses (Ahumada & Lovell, 1971; Beard & Ahumada, 1998). The resulting map of correlations shows the influence of each noise pixel on the observer’s responses, and is called a *classification image*. The classification image reveals which stimulus regions the observer used to perform the task, and can be thought of as a behavioral receptive field.

Gold, Murray, Bennett, and Sekuler (2000) used the response classification method to investigate observers’ strategies in the task developed by Ringach and Shapley (1996). The middle panel of Figure 7 shows the classification images that Gold *et al.* obtained in the real, occluded, and fragmented conditions. Gold *et al.* found that in the fragmented task observers used one or two edges of a single inducer, whereas in the occluded task observers used whole sides of the amodally completed square, including both inducer edges and the physically empty regions between the inducers. Furthermore, Gold *et al.* found that classification images in the occluded condition were similar to classification images in the real condition. The similarity between the observers’ strategy in the occluded and real conditions suggests that occluded, amodally completed contours are perceptually on a par with real, luminance-defined contours, in the sense that observers adopt the same strategy in a perceptual task regardless of whether the stimuli to be discriminated are defined by amodally completed contours or by luminance contours.

This is a surprising result, as it shows that in the occluded condition observers allowed their decisions to be affected by large regions between the inducers that contained no information as to the correct response, which is a strategy one might expect to be quite inefficient. Yet, both Ringach and Shapley (Ringach & Shapley, 1996) and Murray *et al.* (2000) found that performance was *better* in the occluded condition than in the fragmented condition. This paradox is resolved by Murray *et al.*’s finding that, despite appearances, the classification images in the occluded condition do in fact reflect a more efficient strategy than the classification images in the fragmented condition. They showed this by cross-correlating the observers’
classification images with the classification image that would be obtained from the ideal observer for this task (i.e., a hypothetical observer that makes optimal use of all of the available stimulus information). This analysis showed that the cross-correlation of the occluded classification image with the ideal occluded classification image was in fact higher than the cross-correlation of the fragmented classification image with the ideal fragmented classification image. This result implies that observers made more efficient use of the informative stimulus regions (i.e., the luminance-defined inducer edges) in the occluded condition than in the fragmented condition. One can get a sense of this difference by comparing the classification images for occluded and fragmented conditions (Figure 7). Whereas observers used information from only one inducer in the fragmented condition, observers used information from several inducers in the occluded condition.

We suggest that amodal completion is not an end in itself, but a means to a greater end. Amodal completion shifts observers' strategies through grouping, so that spatially disparate fragments of a stimulus are processed as a perceptual unit, thus improving an observer's ability to extract relevant information from a stimulus.

**SUMMARY**

Amodal completion is a critical component of object recognition. Research over the past few decades has provided convincing objective evidence that the visual system does in fact treat partly occluded objects as though they are complete, and the results from these studies can be taken as a case study for the more general problem of grouping. These studies have shown that biologically plausible, computational models of grouping are constrained by several empirical facts concerning the completion of partly occluded objects.

1. Amodal completion takes time. Several different paradigms have been used to objectively measure the time course of completion.

2. Time-to-completion depends on context. Time-to-completion is influenced by factors such as the amount of occlusion, and the presence of consistent stereoscopic depth cues.

3. Strength of completion depends on context. One can derive an objective measure of the strength of completion, and this measure can vary independently of time-to-completion. For
example, in some cases, the addition of motion cues may increase the strength of completion without affecting time-to-completion.

4. The form of completion depends on context. Both the spatial structure of an object and past perceptual experience can influence the form of completion. Local and global processes can both play a role in completion, but their relative weights depend on the context. Under some circumstances, past experience with a particular object may also influence whether and how that object is completed.

5. Completion affects processing efficiency. Noise masking studies have shown that amodal completion affects the efficiency of visual processing without affecting the level of internal noise. The increased efficiency is reflected in a change in the behavioral receptive field revealed by the response classification method. For examples, observers sometimes use more of an amodally completed stimulus than of a very similar fragmented stimulus. This suggests that completion is not as an end in itself, but a means toward an end: Amodal completion, and grouping more generally, enables the observer to use stimuli more efficiently to perform perceptual tasks.

REFERENCES


