Available online at www.sciencedirect.com
SCIENCE DIRECT®

Vision Research 45 (2005) 3262–3276

Vision Research

www.elsevier.com/locate/visres

The recognition of partially visible natural objects in the presence and absence of their occluders

Jeffrey S. Johnson [a,*], Bruno A. Olshausen [a,b]

[a] *Center for Neuroscience, University of California Davis, Davis, CA, USA*
[b] *Redwood Neuroscience Institute, Menlo Park, CA, USA*

Received 21 December 2004; received in revised form 13 May 2005

Abstract

The visual system is adept at compensating for the missing information in scenes that results from occlusion, but how this is done is not fully understood. In particular, the role of the occluding object in visual processing and its effect on the subsequent recognition of the occluded object is unclear. We report three human behavioral experiments suggesting that the recognition of partially visible objects is facilitated when the missing object information is replaced by an occluder rather than simply removed. Furthermore, we provide EEG evidence suggesting that the processes responsible for facilitated recognition occur relatively early in the visual stream.

Keywords: Object recognition; Occlusion; Amodal completion; Deletion; Natural images; ERP

1. Introduction

The world that we live in is a cluttered one. In contrast to the controlled realm of the laboratory, it is only the rare object which is seen in isolation in our daily lives. Many objects are partially occluded by other intervening objects. Despite this, we do not have the impression as we view the world that it is filled with object fragments—the objects that we see appear to be complete ones. Although our impressions of object wholeness could conceivably arise at a purely conceptual level, there are strong ecological reasons to think that our visual system should understand the natural rules of occlusion and have developed some mechanisms at the perceptual level to account for occlusion and other forms of missing information in a scene in the early or intermediate stages of visual processing (Nakayama, He, & Shimojo, 1995).

There are two main types of completion effects that compensate for missing or ambiguous information in the retinal image: *modal* and *amodal* completion (Michotte, Thinès, & Crabbé, 1964/1991). Modal completion is a process that results in effects such as illusory contours (Kanizsa, 1979) and neon color spreading (van Tuijl, 1975). Modal completion is perceptually salient despite having no physical counterpart in the retinal image. Neural correlates of modal completion have been demonstrated in V2 (von der Heydt, Peterhans, & Baumgartner, 1984) and recently, in V1 (Lee, 2003). Amodal completion is the term used to describe the continuation of object contours and surfaces behind occluders, a process which does not manifest a perceptual counterpart. Because occlusion events are common and the illusory conjunction of unrelated contours is rare, amodal completion is more applicable to natural images than modal completion. There are suggestions that the two types of completion are mediated by the same mechanisms (Kellman, Yin, & Shipley, 1998, but see Singh, 2004), but the neural correlates of amodal completion that have been seen are weak in comparison

* Corresponding author. Tel.: +1 530 757 8789.
E-mail address: jsjohnson@ucdavis-alumni.com (J.S. Johnson).

0042-6989/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2005.06.007

to those arising from modal completion (Lee, 2003; Sugita, 1999), and the fact that amodal effects do not result in visible contours suggests that they may be post-perceptual.

Despite the ubiquity of occlusion in the world, most models of visual object recognition are not specifically equipped to account for amodal completion effects, instead focusing their efforts on the goodness-of-match of a feedforward analysis of the image with an object model (Fukushima, 1980; Mel, 1997; Riesenhuber & Poggio, 1999; Ullman & Bart, 2004; Ullman, Vidal-Naquet, & Sali, 2002; VanRullen & Thorpe, 2002). When object fragments are missing, these models do not perform completion, relying instead on matching only the present fragments to the object model. Such an approach is relatively easy to implement, but does not take into account depth-based image cues which distinguish between objects which are partially visible due to occlusion and objects which are partially visible because some of the object is missing. Similarly, a model that performs completion indiscriminately (e.g., Kellman, Guttman, & Wickens, 2001) cannot distinguish between these two cases without resorting to higher-level information to prune inappropriate completions after they have been made (Kellman, 2003). Models such as these predict that recognition (or at least early visual processing) of partially visible objects will not be affected by the presence or absence of an occluder.

On the other hand, a model such as Biederman's Recognition-By-Components (Biederman, 1987; Hummel & Biederman, 1992) makes an explicit attempt to determine which edges in the scene should be bound together using local contour junction rules. Going a step further, some models (Fukushima, 2005; Lee & Mumford, 2003; Nakayama et al., 1995) begin by establishing, with the help of feedback, a global surface-based representation of the scene at low levels of the system. Such models employ inferred depth relations from the earliest representations of the scene and explicitly predict that the visual system will treat image fragments as a single object under occluded conditions (when amodal completion should occur), but separately when global image structure suggests completion is not appropriate. These models suggest that the presence or absence of a depth-appropriate occluder plays a crucial role in determining whether completion occurs, thereby having an effect on the recognition of a partially visible object.

Does the visual system take these depth relations into account when performing amodal completion? One method to determine the effects, if any, that the presence of an occluder has on the recognition of partially visible objects is to construct two sets of images, one with occluded objects and another containing the same object fragments with the occluder removed and the previously occluded regions open to the background. This latter form of image, which by virtue of its depth relations is amodal-inappropriate, we will call "deleted". An early demonstration by Bregman (1981) suggested that a set of outline letterforms which is partially obscured by an occluder in two dimensions is subjectively easier to perceive than the same letter fragments with the occluder removed. Psychophysical studies on similar stimuli have suggested the opposite (Brown & Koch, 2000; Brown & Koch, 1993), showing that subjects are, in general, slower to identify occluded letter fragments than deleted ones. Another study (Gerbino & Salmaso, 1987) using a matching task with outline shape stimuli determined that subjects were faster and more accurate in matching an intact template shape to an occluded version of the shape than to a deleted one. These conflicting studies do not leave a clear picture of the relative difficulty of recognition of occluded and deleted line objects and invite questions about to what extent the visual system entertains depth relations when performing amodal completion.

Another open question is at what stage of the visual pathway the completion of occluded objects is performed. The idea that it may be accomplished quite early is supported by psychophysical studies in visual search of occluded objects (Rensink & Enns, 1998). Neurophysiological evidence of amodal contour responses about 100 ms after presentation of occluded images in macaque V1 cells (Lee, 2003) also supports an early view. However, because occlusion is a function of the relative depth of the objects in the scene, the ability to distinguish an occluded object from a deleted one would appear to rely on an initial determination of a scene's depth relations. Peterson and Gibson (1994) have found behavioral evidence that the determination of depth relations can be contingent upon object contour cues, suggesting that depth relations may be assigned relatively late in visual processing. Furthermore, human event-related potential (ERP) studies on contour closure have suggested that the processes involved in forming a unified percept of a deleted line object are measured no earlier than 230 ms after presentation on electrodes over occipital cortex (Doniger et al., 2000).

The vast majority of the above evidence regarding the efficacy and timecourse of the visual processing of partially visible objects has come from studies of simple shapes and line objects. A notable exception is Nakayama, Shimojo, and Silverman (1989), which shows that photographic face fragments interrupted by bars are easier to recognize when the bars are stereoscopically occluding the face than when the same bars are presented behind the face in a deleted fashion. Such an effect could well rely on stereoscopic depth cues and prove non-replicable in 2D images, even if the relative depth of the objects can be inferred. The studies described here are dedicated to investigating the role of occluders in the recognition of 2D partially visible natural objects. Is