Vision Research 49 (2009) 2475–2484

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

The influence of orientation jitter and motion on contour saliency and object identification

Geir Eliassen Nygård, Tina Van Looy, Johan Wagemans *

Laboratory of Experimental Psychology, University of Leuven, Belgium

ARTICLE INFO

Article history:
Received 6 December 2008
Received in revised form 22 July 2009

Keywords:
Perceptual grouping
Figure-ground segregation
Shape-from-motion
Gabor
Good continuation
Motion perception
Shape perception
Object perception

ABSTRACT

One of the ultimate goals of vision research is to understand how some elements are grouped together and differentiated from others to form object representations in a complex visual scene. There exists an extensive literature on this grouping/segmentation problem, but most of the studies have used un-recognizable stimuli that have little to do with object recognition per se. We used Gabor-rendered outlines of real-world objects to study some relationships between bottom-up and top-down processes in both spatial- and motion form perception. We manipulated low-level properties, such as element orientation and local motion, while incorporating higher-level properties, such as object complexity and identity, and found that adding local motion improved overall performance in both object detection and object identification tasks. Adding orientation jitter effectively decreased object detection performance in both static and motion conditions, and increased reaction time for identification in the static condition. Orientation jitter had much less effect on reaction times for identification in the local motion condition than in the static condition. Both contour properties ("good continuation") and object properties (identifiability) had a positive effect on detection and reaction time for identification.

1. Introduction

A fundamental goal of vision is to locate, characterize, and recognize objects. To determine "what" is "where", the visual system must first determine which parts of the image belong together in groups that can be segmented from the background. This is known as the image grouping/segmentation problem, and one of the most important tasks is the extraction of object contours. Since contours are often not well defined along all of their extent (due to partial occlusion), the visual system needs to be able to infer their nature from an incomplete representation. It can make use of several cues to construct a coherent percept, for instance, texture gradients, color, depth information, occlusion, and motion. How the brain combines local information into a global structure, how it computes form from these cues, remains an important issue in visual neuroscience. Most studies trying to address this subject have used simple, non-object stimuli in detection or discrimination tasks, where identification of objects was not necessarily needed. In an attempt to reduce the gap between the extensive literature on grouping/segmentation and object perception (where object identification has to take place), we introduce a new set of stimuli where we use Gabor-rendered outlines of real-world objects. We will first review some of the background on contour integration and form perception, and then present two experiments that explore the influence of jitter and motion on both the detectability and identifiability of real-world object contours.

Several studies have looked at contour integration in both static and motion conditions. In a paper on the role of temporal modulation in visual contour integration, Bex, Simmers, and Dakin (2001) compared the detectability of "snakes" and "ladders" in relation to orientation jitter on the contour elements. "Snakes" are constituted by Gabor elements that are locally aligned with the contour axis, while "ladders" are constituted by elements perpendicular to its axis. For a contour containing six elements placed along a path with an angle of 20° between the segments, they found that the amount of jitter that could be tolerated was approximately doubled when they added motion to the Gabor elements (translation of the carrier sine wave). In a later study, Bex, Simmers, and Dakin (2003) used a different paradigm (where the contour elements consisted of moving dots) and found that, as with static contour images, the visibility of moving contours decreases at high curvature, albeit by less than in the static case. It has also been found that counter-phase temporal flicker can enhance contour detection (Bex et al., 2001). In exploring the influence of spatial frequency and orientation on motion-defined contours, Ledgeway and Hess (2006) found a very broad tuning for spatial frequency, and a relatively narrow tuning for orientation. Motion direction tuning is in comparison relatively broadband (Allman, Miezin, & McGuinness,

* Corresponding author. Address: University of Leuven, Department of Psychology, Tiensestraat 102, B-3000 Leuven, Belgium. Fax: +32 16 32 60 99.
E-mail address: johan.wagemans@psy.kuleuven.be (J. Wagemans).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.08.002

1985; Bex et al., 2001; Bex et al., 2003; Ledgeway & Hess, 2002, 2006).

Traditionally, mechanisms involved in object recognition and mechanisms for encoding object position and motion have been assumed to project ventrally and dorsally from primary visual cortex to infero-temporal and posterior-parietal cortex, respectively (DeYoe & Van Essen, 1988; Livingstone & Hubel, 1987; Ungerleider & Desimone, 1982), constituting the so-called "ventral" and "dorsal" streams, respectively. Such a strong functional distinction between these two pathways is disputed however: several cortical areas in both the ventral and dorsal streams have been implicated in shape-from-motion (Braddick, O'Brien, Wattam-Bell, Atkinson, & Turner, 2000; Murray, Olshausen, & Woods, 2003) and MT/MST, a typical dorsal area dedicated to motion processing, also appears involved in the analysis of object shape (Kourtzi, Bülthoff, Erb, & Grodd, 2002). Moreover, several studies point to a third stream projecting from V1 to lateral occipito-temporal cortex (LOC) that also underlies complex motion perception (Ferber, Humphrey, & Vilis, 2003, 2005; Grill-Spector, Kushnir, Edelman, Itzchak, & Malach, 1998; Murray et al., 2003). Evidence suggests that the streams consist of a hierarchy of processing stages that transform lower-order stimulus properties into higher-order primitives (Grill-Spector et al., 1998), and anatomical work has revealed reciprocal inter-stream connections at all levels of the visual hierarchy (Felleman & Van Essen, 1991; Van Essen & Maunsell, 1983).

Several studies have set out to find how spatial form and motion form, presumed to be implemented by two independent systems (Ledgeway & Hess, 2006; Lorenceau & Alais, 2001; Rainville & Wilson, 2004), interact when they are presented simultaneously. Lorenceau and Alais (2001), for instance, showed that binding local motions into global object motion depends on spatial form (open vs. closed contour configurations). They suggested that this influence arises in early cortical levels where a spatial-form-based veto of motion integration occurs in the absence of closure. Rainville and Wilson (2004, 2005), on the other hand, argued that the interference is a result of processes further up in the hierarchy involving curvature extraction or overall shape.

The stimuli that have been used in most of the above studies have been relatively simplistic in nature, consisting of geometric figures (squares, circles, polygons), parametric contours (radial frequency patterns), snake-like contour segments, or dot patterns, with little, if any, biological significance. These studies are invaluable because they allow parametric control on all low-level aspects, while restricting the variability and complexity of the shape properties. The downside is, however, that they do not allow to extrapolate the findings to more complex, natural shapes and to link the grouping/segmentation processes to higher-order processes such as object recognition. In an attempt to extend the topics of the ongoing discussion to a domain that is more directly related to object perception, we introduce a new set of stimuli where we use Gabor-rendered outlines of real-world objects.[1] This will enable us to manipulate low-level properties that can be used in models of contour perception, while incorporating higher-level object properties such as complexity and identity. By adding motion we can then study both spatial form and motion form using more familiar stimuli.

Of course, we do not mean to suggest that one cannot study high-level shape and contour processing with synthetically generated stimuli that are not derived from real-world objects. On the contrary, quite interesting shape perception work has made use of synthetic shapes consisting of combinations of Fourier dimensions (e.g., Cortese & Dyre, 1996), well-controlled contours generated by radial frequency patterns (e.g., Bell, Badcock, Wilson, & Wilkinson, 2007; Wilkinson, Wilson, & Habak, 1998), or even synthetic faces (e.g., Wilson, Loffler, & Wilkinson, 2002). Indeed, we have made use of similar shapes with variable levels of complexity and symmetry in our own work (e.g., Kayaert & Wagemans, 2009; Machilsen, Pauwels, & Wagemans, in press; Op de Beeck, Wagemans, & Vogels, 2003). What we do mean to suggest is that stimuli derived from real-world objects probably also induce some extra processing in the highest levels in the visual hierarchy where contact is made with representations of existing objects and associations with other items in semantic memory become available too. Investigation of how good continuity, orientation jitter, local motion, etc. affect the visual processing at these highest levels of the visual hierarchy, and vice versa, requires the kind of stimuli we introduce here.

In sum, such stimuli will enable us to answer one of the original questions of Gestalt psychology: does grouping help identification when the contour represents a familiar object (Wertheimer, 1938)? And vice versa? In Experiment 1, we examine the role of static and dynamic grouping in the detection of these complex shapes and we include good continuation and identifiability as additional variables of interest. In Experiment 2, we explicitly ask to what extent static and dynamic grouping influence the identification of these Gaborized outlines. In addition to addressing the two-way linkage between grouping and identification, we use these stimuli to be able to generalize the earlier findings obtained with simpler, parametrically controlled stimuli to more complex, natural shapes. In a similar way as for snake detection paradigms, we expect objects with more curved segments to be more difficult to both detect and identify, and that adding orientation jitter will degrade performance in both tasks. Furthermore, we expect motion to enhance both detection and identification of our object contours. We also expect to see an influence of global object properties, for instance, a positive effect for objects that are easier to identify.

2. Experiment 1

In Experiment 1 we asked for detection of objects in a noisy background, defined by contours with variable jitter levels on the orientation of the constituent elements. The Gabor elements were either static or had "local motion", i.e., the phase of the Gabors was translated as to give a motion direction perpendicular to the element's orientation while the Gaussian envelope remained stationary. The objective was to see if and how jitter and motion interact in determining detection.

2.1. Subjects

Subjects ($N = 6$) were three male and three female, aged 20–24, with normal or corrected-to-normal vision. One subject is the second author, the remaining subjects were recruited in the general student population for a paid (per hour) participation, and were naïve regarding the purpose and the details of the experiment.

2.2. Stimuli

The stimuli consisted of Gabor elements that were placed and oriented such that they gave rise to the percept of an object embedded in a background (Fig. 1). The objects were contour versions of 60 items from the Snodgrass and Vanderwart (1980) set of line drawings, which we had first converted into silhouettes and then into outlines (De Winter & Wagemans, 2004; Wagemans et al., 2008). Of course, contours of real-world objects differ on a

[1] We have also performed a parallel series of experiments where we used kinetic dot versions of the same objects/stimuli to look at related questions (see Segaert, Nygård, & Wagemans, 2009).