Available online at www.sciencedirect.com
ScienceDirect
Vision Research 47 (2007) 2825–2840
Vision Research
www.elsevier.com/locate/visres

Geometric determinants of shape segmentation: Tests using segment identification

Elias H. Cohen *, Manish Singh

Department of Psychology and Center for Cognitive Science, Rutgers University, New Brunswick, USA

Received 16 February 2007; received in revised form 6 June 2007

Abstract

The geometric determinants of shape decomposition were studied using a performance-based method. Observers' identification of contour segments was shown to be systematically modulated by their curvature properties, and by the geometric properties of the enclosed region. Specifically, negative minima of contour curvature provided the best segment boundaries. Segments with negative-minima boundaries were identified with greater accuracy than those with positive maxima or inflection boundaries of comparable length. Additionally, segment identification was shown to be determined by contour length, the turning angle at part boundaries, and the width at the part's base (hence the part's protrusion). The results indicate that part decomposition is an automatic process. Moreover, this process is graded, i.e. parts are more strongly segmented, or more likely to be perceived, according to the strength of many geometric determinants.

Keywords: Shape representation; Part segmentation; Parts; Contour curvature; Structural descriptions; Part salience

1. Introduction

Understanding of the neural basis of visual shape has, for many years, been largely restricted to very early or very late stages of processing. Early visual cortical areas (such as area V1) are known to respond to simple local image features, such as oriented edges (Hubel & Wiesel, 1959, 1962). Work on late visual processing (such as area IT), on the other hand, has shown selectivity for high-level categories of complex recognizable shapes, such as faces or hands (Desimone, Albright, Gross, & Bruce, 1984; Gross, Rocha-Miranda, & Bender, 1972; Perrett, Rolls, & Caan, 1982; Tanaka, Saito, Fukada, & Moriya, 1991). However, processing at intermediate levels, responsible for transforming local image measurements into global representations of objects, is only beginning to be understood (Pasupathy & Connor, 1999, 2001, 2002). It is at these intermediate-level representations of shape that we focus our psychophysical investigation. Despite a great deal of discussion of parts in the shape literature, surprisingly few performance-based methods have been developed to investigate the part-based nature of visual shape. This paper reports a series of psychophysical experiments designed to quantify geometric influences on the part-based representation of visual shape. In doing so, it provides some indication of how the gap between local measurements of edges and the global representation of shape might be bridged.

1.1. Shape and part

A great deal of information about an object's shape is carried by its *occluding contour* (Attneave, 1954; Koenderink, 1984). Indeed, recognition performance with silhouettes has sometimes been found to be almost as good as with shaded drawings of 3D objects (Hayward, Tarr, &

* Corresponding author. Address: Department of Vision Sciences, State University of New York, State College of Optometry, 33 West 42nd Street, New York, NY 10036, USA.
E-mail address: ecohen@sunyopt.edu (E.H. Cohen).

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.06.021

Corderoy, 1999). Recent work using single-cell recordings in area V4 in the primate visual cortex has investigated the representation of this important source of information, and found evidence for a *piecewise* coding of shape. Specifically, Pasupathy and Connor (1999) demonstrated a selectivity of V4 cells to contour curvature as determined by *curvature polarity* (whether a contour segment has positive or negative curvature—i.e., bounds a locally convex or concave region) and *curvature magnitude* (or contour "acuteness"). Different populations of cells were found to respond selectively to either convex or concave segments of an occluding contour. In subsequent studies employing whole shapes, Pasupathy and Connor (2001, 2002) found evidence for a distributed coding of shape, with different units within a population responding to either convex or concave segments at different angular positions on the shape (relative to its center). The representation of shape in V4 thus appears to be a *piecewise* representation, in that different units in V4 encode different—either convex or concave—pieces or segments of a shape's occluding contour.

The notion of a part-based representation has also played an important role in object recognition and behavioral approaches to visual shape. Although any subset of a shape may, in a sense, be considered its "part," the perceptual notion of a part corresponds to those subsets of a shape that are naturally perceived as being semi-independent units—hence perceptually separable from the rest of that shape. Parts are thus more than the arbitrary pooling of earlier stage inputs; they are rather the perceptually and psychologically salient visual units of a shape's representation. Organizing shape representations in terms of parts—with smaller parts hierarchically nested within larger parts—allows one to separate the representation of the shape of each individual part from the representation of the spatial relationships between the parts. This, in turn, leads to a more robust representation of shape—one that is more stable across viewing conditions—e.g., changes in the articulated pose of an object, or an observer's vantage point with respect to it (Biederman, 1987; Hoffman & Richards, 1984; Marr & Nishihara, 1978; Palmer, 1977). Furthermore, hierarchical structure in shape representation is a central aspect of visual experience. The ability to direct action to different levels of object structure is crucial to visually guided manipulation and interaction. People, arms, hands, and fingers are all natural candidates for visual attention, semantic labeling, and motor interaction. Understanding the geometric determinants of shape segmentation thus allows one to understand how these separable units of shape are perceptually generated.

Part-based representation of shape, while central to many theories of high-level vision and object recognition, has generally lacked systematic psychophysical investigation. Compelling psychophysical support for part-based representation of shape should ideally take the following form: (i) demonstration that human vision automatically divides complex shapes into smaller units, (ii) demonstration that there are consistent and predictable rules that dictate the segmentation of shape into these sub-units, and that these sub-units determine the storage of visual information, and finally, (iii) demonstration that these same sub-units dictate the extraction of other visual properties, such as orientation, location, and size. The following investigations were undertaken with points (i) and (ii) as the guiding framework. (For work investigating point (iii), see Cohen & Singh, 2006; Denisova, Singh, & Kowler, 2006).

Parts have generally been studied from the point of view of high-level vision, as categorical units of object representation (Tversky & Hemenway, 1984) and units important for recognition and naming (e.g., Biederman, 1987; Biederman & Cooper, 1991). However, what a part *is* from the point of view of bottom–up visual processing (i.e., low-level mechanisms of visual segmentation) is less clear. Our goal is to characterize the notion of a visual part in concrete psychophysical terms, based on observers' performance in an objectively-defined task. Unlike many previous studies, we use unfamiliar randomly-generated shapes, in order to focus on the geometric properties of the bounding contour. In a part-based account, certain portions of a shape constitute natural units of representation for the visual system, and therefore should be much more readily identifiable than other portions of comparable size. We therefore use segment identification as the operational test for naturalness of shape parts. The "null hypothesis" then becomes identification performance as might be predicted by a decomposition-free account of shape representation (such as one based on unstructured templates). Under such an hypothesis, any portion of a shape should be equally easy or difficult to identify, as long as its size is preserved.

In this paper, we focus on one source of shape information, namely occluding contour. Specifically, we examine the geometric properties of a contour segment that make it more or less identifiable. Experiments 1 and 2 compare contour curvature landmarks that potentially provide natural *boundary cues* for parts. Experiments 3 and 4 examine geometric factors that determine the *salience* of a part's representation.

1.2. Contour curvature and information content

Attneave (1954) observed that objects' boundaries—their occluding contours in the projected image—have high information content as they signify the greatest change in image characteristics. Koenderink (1984) has shown, more specifically, that the occluding contour carries a great deal of information about 3D shape: for smooth shapes, the sign of curvature of the occluding contour directly informs one of the sign of Gaussian curvature of the 3D shape.

Similarly, along occluding contours, Attneave (1954) observed that information is concentrated at extrema of contour curvature—points where the change signified by curvature is the highest (i.e. points where the magnitude