Vision Research 126 (2016) 347–361

Contents lists available at ScienceDirect

Vision Research

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

Investigating shape representation using sensitivity to part- and axis-based transformations

Kristina Denisova *, Jacob Feldman, Xiaotao Su, Manish Singh

Department of Psychology and Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ, United States

ARTICLE INFO

Article history:
Received 4 February 2015
Received in revised form 6 July 2015
Accepted 10 July 2015
Available online 28 January 2016

Keywords:
Shape
Parts
Axes
Shape skeleton
Non-rigid transformations
Shape discrimination

ABSTRACT

Part- and axis-based approaches organize shape representations in terms of simple parts and their spatial relationships. Shape transformations that alter qualitative part structure have been shown to be more detectable than those that preserve it. We compared sensitivity to various transformations that change *quantitative* properties of parts and their spatial relationships, while preserving qualitative part structure. Shape transformations involving changes in length, width, curvature, orientation and location were applied to a small part attached to a larger base of a two-part shape. Increment thresholds were estimated for each transformation using a 2IFC procedure. Thresholds were converted into common units of shape difference to enable comparisons across transformations. Higher sensitivity was consistently found for transformations involving a parameter of a single part (length, width, curvature) than those involving spatial relations between two parts (relative orientation and location), suggesting a single-part superiority effect. Moreover, sensitivity to shifts in part location – a biomechanically implausible shape transformation – was consistently poorest. The influence of region-based geometry was investigated via stereoscopic manipulation of figure and ground. Sensitivity was compared across positive parts (protrusions) and negative parts (indentations) for transformations involving a change in orientation or location. For changes in part orientation (biomechanically plausible), sensitivity was better for positive than negative parts; whereas for changes in part location (biomechanically implausible), no systematic difference was observed.

1. Introduction

A fundamental question in visual perception is how the human visual system represents object shape. Part of the difficulty in addressing this question is that the notion of shape itself is not as clear-cut as many other visual attributes. One definition of shape, originating in geometry and mathematics, is that it refers to geometric properties that are unaffected by rigid transformations (transformations that preserve inter-point distances) and uniform scaling. This is consistent with intuition: moving a statue to a different location, making it face in a different direction, or even having a replica made that is half the original size, does not alter what we consider to be “its shape”. Under this view, two shapes are equivalent if they can be brought into alignment by applying one or more of these transformations (e.g. Ullman, 1989). As a result, the mathematical “distance” between two shapes can be defined in terms of how much they still differ once they have been brought into maximal alignment using these transformations (e.g. Kendall, 1989; Mardia & Dryden, 1998).

Psychologically speaking things are more complicated, however, and the above definition fails to capture the notion of “perceived shape”– i.e. as far as representation by the human visual system is concerned. First, two forms that are geometrically equivalent can look different. Fig. 1A shows an example due to Mach (1897/1914/1959) in which the two shapes look different (“square” versus “diamond”) despite the fact that they differ only by a rigid rotation in the image plane through 45°. Indeed, there is a great deal of work demonstrating large effects of orientation on shape perception (e.g., Rock, 1973; Tarr & Pinker, 1989). Second, geometrically distinct shapes (i.e. shapes not related by rigid transformations and uniform scaling, or even by affine transformations) are often perceived to be the same. Consider a cat viewed in a crouching position versus one in a pouncing pose (Fig. 1B). Technically, the two shapes are of course different: no rigid template could possibly work for both shapes. Yet there is a sense in which the two shapes

* Corresponding author at: Department of Psychiatry, College of Physicians and Surgeons, Columbia University, 1051 Riverside Drive, Unit 40, New York, NY 10033, United States.

E-mail addresses: denisova@nyspi.columbia.edu (K. Denisova), jacob@ruccs.rutgers.edu (J. Feldman), xsu@ruccs.rutgers.edu (X. Su), manish@ruccs.rutgers.edu (M. Singh).

http://dx.doi.org/10.1016/j.visres.2015.07.004
0042-6989/

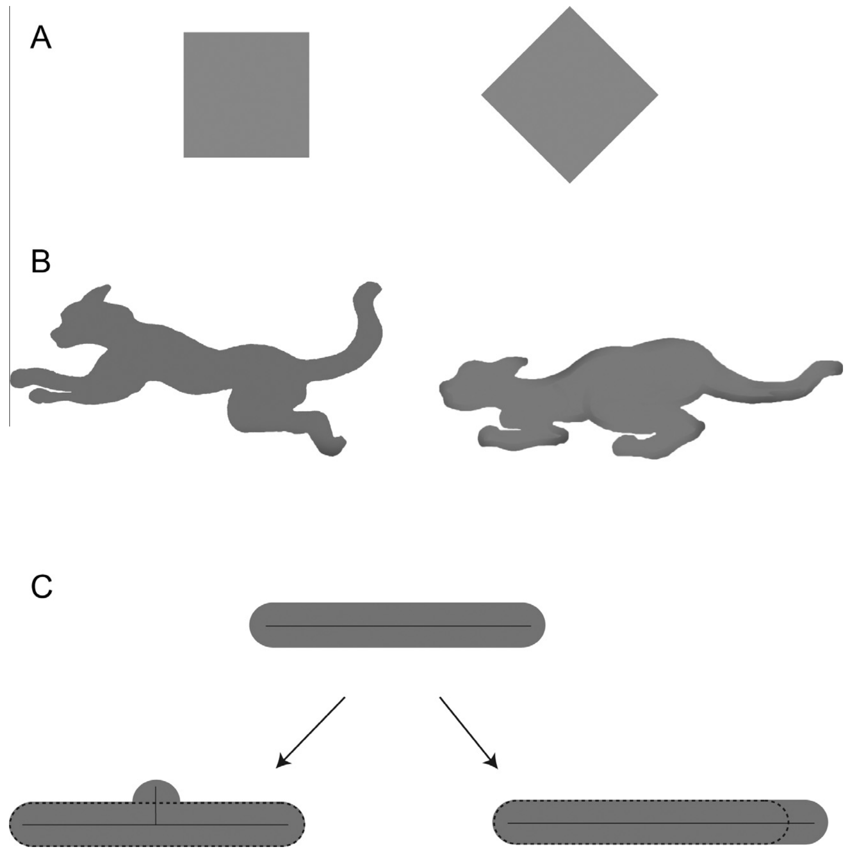

Fig. 1. Clarifying the inadequacy of geometric equivalence in capturing shape perception. (A) Two geometrically equivalent shapes that look different: "square" vs. "diamond" (Figure adapted from Mach, 1897/1914/1959). (B) Two geometrically distinct shapes (i.e., they cannot be related by rigid transformations and uniform scalings) that look perceptually similar. These two are naturally perceived as arising from actions of the same biological organism. The cat is shown in two different poses, each with different configurations of its limbs. (C) Shape transformations that alter qualitative part structure are perceived as larger changes ("bump" on the bottom left) than those that preserve part structure ("extension" on the bottom right). Note that the transformation on the right actually involves a larger physical change to the original (top) shape.

are in fact equivalent, namely, the two differ only in the articulation of the limbs of the same biological form. Such part articulations are in fact extremely common in animate objects, and define an important class of shape transformations that the visual system must deal with successfully.

1.1. Part-based representation of shape

The non-rigid movements of biological forms have been an important source of motivation for the part-based approach to shape representation. According the part-based approach, the visual system represents the shape of complex objects in terms of simpler parts, and the spatial relationships between these parts. In other words, it proceeds by segmenting complex shapes into parts, and organizing the shape representation as a hierarchy of parts. An important feature of this "structural" approach to shape is that it separates the representation of the shape of the individual parts from the representation of the spatial relationships between these parts. As a result, an object can be readily identified as composed of the same parts as another object, though in slightly different (but still "valid") spatial relationships. This property allows part-based representations to be more robust to changes in the articulated pose of an object: the cat can be recognized as being essentially the same "form," irrespective of whether it is sleeping or running.

An important cue for segmenting a shape into parts is the presence of negative minima of curvature along its bounding contour (Hoffman & Richards, 1984). Negative minima of curvature are points with locally maximal magnitude of curvature that lie in concave regions of shape (i.e. regions with negative curvature). The minima rule is motivated by a regularity of nature known as transversality: the process of joining two separate objects to form a single composite object generically produces negative minima – in this case, tangent discontinuities – at their locus of intersection (Hoffman & Richards, 1984). Similarly, the sprouting of a new part (say, from a seed or embryo) similarly produces negative minima of curvature (Leyton, 1989). Thus, given only a composite shape, points of negative minima provide natural candidate points for segmenting the shape into parts. This is not to say that negative minima are sufficient by themselves to divide shapes into parts. First, they do not specify how candidate boundary points should be paired to form cuts that divide the shape. Second, negative minima can fail to be part boundaries (e.g. negative minima along a bending snake), and part boundaries can fail to be negative minima (see Barenholtz & Feldman, 2003; De Winter & Wagemans, 2006; Siddiqi, Tresness, & Kimia, 1996; Singh, 2014; Singh & Hoffman, 2001; Singh, Seyranian, & Hoffman, 1999). Negative minima do, however, provide an important cue for part segmentation.

A great deal of psychophysical work has shown that the visual system represents shapes in terms of parts (Biederman, 1987; Biederman & Cooper, 1991; Cave & Kosslyn, 1993; Hayworth & Biederman, 2006; Hoffman & Richards, 1984; Hoffman & Singh, 1997; Lamberts & Freeman, 1999; Lamote & Wagemans, 1999), and that this has important implications for a number of perceptual phenomena, including figure/ground assignment (e.g. Barenholtz & Feldman, 2006; Baylis & Driver, 1995; Hoffman & Singh, 1997; Kim & Feldman, 2009; Stevens & Brookes, 1988), change detection (Barenholtz, Cohen, Feldman, & Singh, 2003; Bertamini & Croucher, 2003; Bertamini & Farrant, 2005; Cohen, Barenholtz,