Vision Research 94 (2014) 33–40

Contents lists available at ScienceDirect

Vision Research

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

Shape constancy measured by a canonical-shape method

Ian P. Howard, Yoshitaka Fujii, Robert S. Allison *, Ramy Kirollos

Centre for Vision Research, York University, Toronto, Ontario M3J 1P3, Canada

ARTICLE INFO

Article history:
Received 18 May 2013
Received in revised form 28 October 2013
Available online 9 November 2013

Keywords:
Shape constancy
Perspective
Binocular disparity

ABSTRACT

Shape constancy is the ability to perceive that a shape remains the same when seen in different orientations. It has usually been measured by asking subjects to match a shape in the frontal plane with an inclined shape. But this method is subject to ambiguity. In Experiment 1 we used a canonical-shape method, which is not subject to ambiguity. Observers selected from a set of inclined trapezoids the one that most resembled a rectangle (the canonical shape). This task requires subjects to register the linear perspective of the image, and the distance and inclination of the stimulus. For inclinations of 30° and 60° and distances up to 1 m, subjects were able to distinguish between a rectangle and a trapezoid tapered 0.4°. As the distance of the stimulus increased to 3 m, linear perspective became increasingly perceived as taper. In Experiment 2 subjects matched the perceived inclination of an inclined rectangle, in which the only cue to inclination was disparity, to the perceived inclination of a rectangle with all depth cues present. As the distance of the stimulus increased, subjects increasingly underestimated the inclination of the rectangle. We show that this pattern of inclination underestimation explains the distance-dependent bias in taper judgments found in Experiment 1.

1. Introduction

When a flat object is rotated out of the frontal plane the shape of its retinal image changes. Simple shape constancy refers to the ability to perceive that a flat object remains the same shape when seen in different orientations relative to the frontal plane. Shape constancy depends on the accurate registration of the shape of the retinal image and of the orientation of the object with respect to the frontal plane (Fig. 1). The orientation, in turn, can be estimated from depth cues such as perspective and stereopsis.

For a shape defined only by its outline, perspective can be divided into linear perspective (where parallel edges or lines extending in depth project to converging lines in a perspective projection) and foreshortening (aspect ratio). Fig. 2a shows the perspective produced by inclining a rectangle about a horizontal axis. In the absence of other depth cues, perspective provides information about inclination[1] only for certain shapes. Linear perspective, but not foreshortening, provides information about inclination for shapes with parallel sides but no specified ratio of width to height, such as a rectangle. In our experiments we were concerned only with linear perspective produced by an outline rectangle.

Binocular disparity (differential perspective) provides information about inclination for all shapes. However, the disparity signal is strong for shapes such as rectangles, and weak for circles due the large amount of vertically-oriented contour in the former from which angular disparity can be reliably extracted. An inclined rectangle produces an angular disparity between the left and right sides, as shown in Fig. 2b. However, angular disparity is inversely proportional to viewing distance and must therefore be scaled by distance. The angle of convergence of the eyes and the associated accommodation provide the only cues to the distance of an outline shape viewed in dark surroundings. Several investigators have reported that people are reasonably accurate at judging the distance of an object within arm's reach when the only cue to distance is the angle of convergence of the eyes. When subjects pointed with an unseen hand to a disc of light at vergence-specified distances of 25, 30, or 40 cm they were accurate to within 1 cm (Swenson, 1932). However, judgments of distance become more variable as distance increases. For example, the variability of setting an unseen rod to the same distance as a binocularly viewed vertical rod increased as vergence-specified distance increased from 30 to 100 cm (Tresilian, Mon-Williams, & Kelly, 1999).

In the absence of all depth information it is not possible to judge the true shape of any object because its image could arise from any combination of shape and orientation that projects that image, as illustrated in Fig. 1. This geometrical relationship between shape and orientation prompted Koffka (1935) to propose that an error in the perceived inclination of a flat object is accompanied by a corresponding error in the judged shape of the object. This is known as the shape-inclination invariance hypothesis.

Shape constancy has been the subject of much investigation and controversy but a comprehensive review is beyond the scope of

* Corresponding author. Address: York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada. Fax: +1 416 736 5872.
E-mail address: allison@cse.yorku.ca (R.S. Allison).

[1] The term *inclination* here refers to rotation of a surface in depth about a horizontal axis (slant with a surface-tilt angle of 90° in Steven's (1983) parameterization).

0042-6989/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2013.10.021

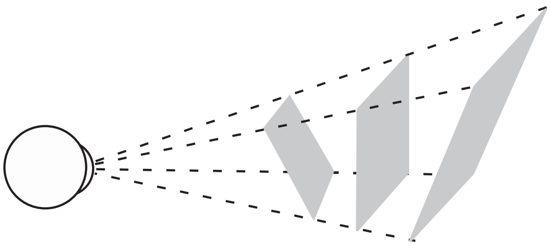

Fig. 1. A frontal rectangle produces the same retinal image as any inclined trapezoid that fills the same visual pyramid. The greater the inclination of the object, the more tapered it must be to produce the same image as the rectangle.

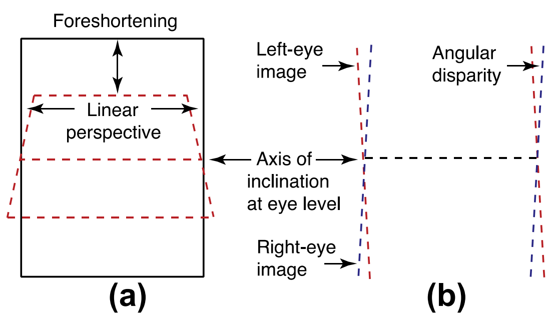

Fig. 2. (a) The linear perspective and foreshortening produced by an inclined rectangle. The dashed shape is the projection of the inclined rectangle onto the frontal plane. The projections of the parallel sides of an inclined rectangle are tapered due to linear perspective and the extensions of these projections intersect at an angle, *p*, which describes the taper due to perspective. (b) The angular disparity produced by the left and right edges of an inclined rectangle.

this paper (for a recent review see Howard, 2012, pp. 137–144). Shape constancy progressively fails as cues for 3-D orientation of the stimulus are weakened (Joynson & Newson, 1962; Thouless, 1931). However, experiments to prove the shape-inclination invariance hypothesis quantitatively have produced variable results because they suffered from procedural problems (Epstein & Park, 1963). One problem is that when asked to judge the shape of a frontal ellipse a person may state that it is a circle because most elliptical images arise from circles. But when asked to judge the orientation of the ellipse they may ignore its shape and concentrate on information that indicates that it is frontal.[2] A second problem is that simple shapes seen in dark surroundings appear to fluctuate over time so that a shape judgment made at one instant may have no relation to a judgment of inclination made at another instant. A third problem is that a person may correctly register the relation between two stimulus features but be unable to register the single features correctly.

Three methods have been used to measure shape constancy. Thouless (1931) described the most frequently used method. Subjects select a shape in the frontal plane to match a shape inclined about a horizontal axis or slanted about a vertical axis. For example, subjects select an ellipse in the frontal plane that matches an inclined circle. Subjects typically select a frontal ellipse that is intermediate between a circle and the elliptical image of the inclined circle. Thouless referred to this as "regression to the real object".

There are two major problems with the Thouless method. Joynson (1958) pointed out that results depend on how subjects interpret the instructions. They could attempt to select a frontal shape that matches the image of the inclined shape. We will call this *image matching*. Otherwise, subjects could attempt to select a frontal shape that matches the actual inclined shape. We will call this *shape matching*. Even when subjects are instructed to respond in a certain way they may fail to follow the instructions (Kaess, 1978). The phrase "regression to the real object" presupposes that subjects are trying to match images. If they were trying to match actual shapes (shape constancy) one would have to say that they "regressed to the image".

The second problem with the Thouless method is that the frontal stimulus may be perceived inaccurately. For example, a frontal ellipse may be perceived as an inclined circle, which produces the impression that the minor axis of the elliptical image is elongated. This may occur even though other information indicates that the ellipse lies on a frontal plane. Thus one cannot know whether a failure of shape constancy arises from inaccurate perception of the inclined stimulus or of the frontal stimulus or of both. Epstein and Park (1963) reviewed these and other methodological problems with the Thouless method.

The second procedure for measuring shape constancy is to ask subjects to draw an inclined shape on a vertical surface. Thouless (1931), Clark, Smith, and Rabe (1956), and Nelson and Bartley (1956) used this method. It is a very unsatisfactory procedure because subjects may attempt to draw in perspective rather than draw the actual shape of the inclined stimulus. In any case, most people draw an object in perspective very inaccurately even though they perceive the shape of the object accurately (Howard & Allison, 2011).

In a third procedure, which we will call the *canonical-shape method*, subjects select from a series of inclined stimuli the one that most resembles a defined shape such as a circle, cross, regular polygon, or rectangle. These are all uniquely defined (canonical) shapes. The method provides a direct measure of shape constancy and avoids ambiguities associated with comparing two shapes at different inclinations. It is highly unlikely that subjects would adopt an image matching approach in this task. While retinal shape could theoretically be judged (that is, is the retinal image a rectangular image?), this is not natural and the subjects are explicitly instructed to report the true physical shape of the object. Furthermore, unlike the Thouless method, there is no comparison with shapes on a frontal surface; shapes that can be readily interpreted as both the physical shape or as a rendering/ projection of a shape on that frontal plane. Two investigators have used the canonical-shape method.

Stavrianos (1945) displayed a set of rectangles simultaneously on an inclined rectangular board viewed binocularly at a distance of 60 cm. The rectangles varied in height and subjects selected the one that most resembled a square. The results indicated a high level of shape constancy. As inclination was increased from 0° (vertical) to 55° the rectangle selected as square increased in height by 6% and the rectangle appeared less inclined. These results are what one would expect from shape-inclination invariance. However, the rectangular board upon which the stimuli were mounted was visible. Stavrianos admitted that this might have produced shape contrast between the board and the shapes on the board.

Saunders and Backus (2006) projected trapezoids one at a time on a frontal screen at a distance of 2 m. They were viewed monocularly. The trapezoids produced the same images as squares inclined to the frontal plane by various degrees about a horizontal axis. The trapezoid judged to be square was considerably greater in height than the image of a square. This indicates that the simulated inclination of the trapezoids was underestimated. This is not surprising because binocular cues to inclination were absent, the

[2] As a reviewer pointed out, it is conceivable that a subject could potentially show the opposite behavior, favoring a frontal interpretation for shape but not orientation judgment. We think this less likely as it deviates from norms (or priors) for both shape and orientation.