Vision Research 126 (2016) 318–329

Contents lists available at ScienceDirect

Vision Research

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

Perception of shape and space across rigid transformations

Filipp Schmidt *, Patrick Spröte, Roland W. Fleming

Justus-Liebig-University Giessen, Germany

ARTICLE INFO

Article history:
Received 30 January 2015
Received in revised form 23 April 2015
Available online 30 April 2015

Keywords:
Perceptual organization
Grouping
Causal history
Transformations
Object constancy
Shape perception

ABSTRACT

Objects in our environment are subject to manifold transformations, either of the physical objects themselves or of the object images on the retina. Despite drastic effects on the objects' physical appearances, we are often able to identify stable objects across transformations and have strong subjective impressions of the transformations themselves. This suggests the brain is equipped with sophisticated mechanisms for inferring both object constancy, and objects' causal history. We employed a dot-matching task to study in geometrical detail the effects of rigid transformations on representations of shape and space. We presented an untransformed 'base shape' on the left side of the screen and its transformed counterpart on the right (rotated, scaled, or both). On each trial, a dot was superimposed at a given location on the contour (Experiment 1) or within and around the shape (Experiment 2). The participant's task was to place a dot at the corresponding location on the right side of the screen. By analyzing correspondence between responses and physical transformations, we tested for object constancy, causal history, and transformation of space. We find that shape representations are remarkably robust against rotation and scaling. Performance is modulated by the type and amount of transformation, as well as by contour saliency. We also find that the representation of space within and around a shape is transformed in line with the shape transformation, as if shape features establish an object-centered reference frame. These findings suggest robust mechanisms for the inference of shape, space and correspondence across transformations.

1. Introduction

A fundamental task of human vision is object recognition. In general, the most important visual feature for object recognition is shape. However, the shapes of objects in our visual environment are subject to manifold transformations, from simple rigid changes like rotation or translation to complex non-rigid transformations like twisting, bending or biological growth. These transformations may be grouped into two broad classes: (i) transformations of the physical objects themselves and (ii) transformations of the object images on the retina. To identify objects across a wide range of viewing conditions, the visual system must be able to represent objects in a way that is robust across transformations from both classes.[1]

Transformations from the first class result from movements and changes of objects, as for example in rolling balls, flying birds, shattering pots, or melting ice cream. In our visual environment movements and changes of objects are often intertwined, for example our limbs undergo non-rigid bending as we walk. Also, many transformations are typical for specific objects so that transformations alone can be powerful cues for object identity (as in biological motion; e.g., Cutting, 1982; Cutting & Kozlowski, 1977; Johansson, 1973). Transformations from the second class result from movements of the observer's eyes, head, and body, as for example in smooth pursuit, turning the head, or movements towards or away from an object.

Many of these transformations have drastic effects on the object's physical appearance and its retinal image. Despite the fact that these transformations occur on a regular basis for objects in our visual environment, we perceive objects that are stable in space and time (*object constancy*; Cassirer, 1944) – as has also been demonstrated for natural images (Kingdom, Field, & Olmos, 2007). Also, we often seem to have some idea about the type of transformation that has given an object its present physical form (*causal history*; Arnheim, 1974, 1988; Leyton, 1989; Pinna, 2010). Both object constancy across transformations and the inference of an object's causal history, are cornerstones of object perception. In this study, we investigate both the perception of object shape

* Corresponding author at: Justus-Liebig-University Giessen, General Psychology, Otto-Behaghel-Str. 10F, 35394 Gießen, Germany. Fax: +49 6419926119.

E-mail address: filipp.schmidt@psychol.uni-giessen.de (F. Schmidt).

[1] This could be achieved by explicitly estimating and then compensating for transformations. For example, if an object is rotated, the visual system could explicitly recover the rotation and discount it. Another possibility is that the visual system effectively 'ignores' the transformations through the use of shape measurements that are invariant across the transformations. For example, it could rely on a variety of measurements, such as the area, or the distance of curvature extrema from the center of mass, which capture certain aspects of shape but which remain constant across rotations.

http://dx.doi.org/10.1016/j.visres.2015.04.011
0042-6989/

(and space), and the inference of transformations across rigid geometric transformations.

1.1. Related work

Before we discuss some relevant findings, we will briefly describe the hierarchy of five groups of transformations that was established by Klein (1893). This hierarchy is defined by the extent to which definitions and theorems of a geometry remain invariant under each group, and has been very useful for describing and integrating findings about transformations in visual perception (Bedford, 2001; Chen, 2005; Graf, 2010). Without any transformation, all properties of a geometry remain invariant (e.g., square → square). In the following, each of the higher-ranking groups includes all transformations of the lower groups. In the first group of Euclidean transformations, properties of size, angle, parallelism, collinearity, order, and connectivity remain invariant – transformations include translations, picture-plane rotations, and reflections (e.g., square → rotated square). In the second group of Similarity transformations, the properties of the first group except size remain invariant – thus, transformations include expansions and extractions (e.g., square → small square). In the third group of Affine transformations, the properties of the second group except angle remain invariant – transformations include stretching and compression of one axis, or rotation of one axis against the other (e.g., square → parallelogram). In the fourth group of Projective transformations, the properties of the third group except parallelism remain invariant – transformations include projection (e.g., square → trapezoid). In the fifth group of Topological transformations, the properties of the fourth group except collinearity remain invariant (i.e., order, connectivity) – transformations include space-curving (e.g., square → circle). Finally, non-topological transformations have no invariant properties (e.g., square → two circles).

Studies on object constancy employed a diverse number of experimental paradigms to measure the effects of numerous kinds of transformation (for reviews see Bedford, 2001; Chen, 2005; Graf, 2010). Here, we will briefly discuss apparent motion paradigms and same-different paradigms together with some standard findings to illustrate typical experimental approaches in the field.

In apparent motion paradigms, sequential presentation of two stationary stimuli at different locations can result in the subjective impression of a single object undergoing motion from one location to the other. As apparent motion mostly occurs when the two stimuli are identical or similar, it is assumed that the two stimuli are interpreted as two glimpses of the same stimulus at two different times (Rock, 1983; Shepard, 1984). Consequently, the apparent motion paradigm can be used to investigate under which conditions observers do perceive two shapes as instances of the same object (e.g., a shape and a transformed version of that shape). Bedford and Mansson (2010) used a competing motion paradigm to directly compare the probabilities with which different transformations induce apparent motion. Thus, the authors tested which transformations are preferred to others by the visual system. They observed that there is a preference to perceive motion between two stimuli that are transformed by Similarity transformations compared to topological transformation (square → small square vs. square → circle/triangle), and a preference to perceive motion between two stimuli that are transformed by topological transformations compared to non-topological transformation (square → circle/triangle vs. square → square with hole). These results suggest that the hierarchy of transformations by Klein (1893) might have an equivalent in visual perception (see also Todd, Weismantel, & Kallie, 2014).

In same-different paradigms, a shape is presented simultaneously with a transformed version of that shape or simultaneously with another (sometimes reflected version of that) shape. Participants have to indicate as quickly as possible whether the two shapes are the same or different when allowing for affine transformations. A general finding is that participants' performance (e.g., response times) decreases with increasing transformational distance between a shape and its transformed version. By far the most studies focused on the effect of rotation, showing that response times increase with increasing angular departure between the two shapes (probably with exception of the principal axes, Lawson & Jolicoeur, 1999; for a review, see Shepard & Cooper, 1982). Also, there is evidence for increasing response times with increasing size ratios (e.g., Bundesen & Larsen, 1975), with increasing translational distance between the two shapes (e.g., Larsen & Bundesen, 1998), or with increasing affine stretching or compression (Dixon & Just, 1978).

In contrast to the effect of transformations on object constancy, there is much less research on the representation of the transformation itself. For most transformed objects, we not only know *that* it was transformed but we have also some idea about *how* it was transformed (i.e., about its causal history; Arnheim, 1974; Leyton, 1989; Pinna, 2010). From a projected object, we might infer its depth plane and from a bent object the type of forces that were applied to it (Spröte & Fleming, 2016). Thus, research on the causal history of objects is interested in the extent to which the current state of an object provides cues for the transformations that operated on it in the past.

Leyton (1989) distinguished between the inference of causal history from a single shape (the 'first inference problem'), and the inference of intervening causal history between a pair of shapes that appear to be snapshots of the same shape at different stages of development (the 'second inference problem'). Inference of causal history from a single shape was investigated by Spröte and Fleming (2013). They tested which geometrical features contribute to our perception of shapes as being "bitten", that is, as having a history of forceful excision of a piece. They found that the relative size and salience of concavities in a shape contribute strongly to our perception of this specific interpretation of the shape's causal history. The inference of causal history between a pair of shapes is somewhat less challenging because different stages can provide information about the dynamics of a transformational process. In fact, if the number of snapshots would be increased sufficiently, participants would be able to closely follow each stage of a transformation. However, even continuous motions can be ambiguous about transformations; for example, a rotating ellipse (or ellipsoid) can be seen as deforming non-rigidly (Jain & Zaidi, 2011; Weiss & Adelson, 2000; Zang, Schrater, & Doerschner, 2010). It remains unclear to what extent participants that are presented with two snapshots can infer the transformation that produced one shape from the other.

A related research question is whether an inferred transformation is restricted to the representation of the transformed shape or whether it extends to the space around the shape. In other words, do participants represent space around a transformed shape in egocentric coordinates (e.g., with respect to their own body) or within a reference frame that is defined relative to that shape? In general, concepts of perceptual reference frames assume that transformations are discounted by imposing a frame of reference that establishes a coordinate system relative to which a shape is perceived (e.g., Rock, 1997). According to this approach, recognition involves the adjustment of a reference frame to the orientation of a stimulus (Graf, Kaping, & Bülthoff, 2005; Jolicoeur, 1990).

Evidence for the existence of object-centered (*allocentric*) reference frames comes from neurophysiology. Some patients with *hemineglect*, a neurological syndrome resulting from a unilateral lesion of the parieto-occipital cortex, show deficits in registering and interacting with the left side of objects or shapes. This left side