

REPLY

Constancy and Similarity

RONEN BASRI*

Department of Applied Mathematics, The Weizmann Institute of Science, Rehovot, 76100, Israel

AND

DAVID JACOBS†

NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540

Received December 12, 1994; accepted November 21, 1995

Pizlo, Rozenfeld, and Weiss define shape constancy as the problem: “... when are shapes perceived as the same?” They propose that two shapes are perceived as the same if they are related by a perspective transformation. We will argue that, on the contrary, human perception of shape constancy cannot be captured by a simple set of geometric transformations. Rather, we claim that the perception that shapes are identical can be influenced by many factors, including the amount of distortion produced by a transformation, the complexity of the shape, and context and prior knowledge. The solution to the problem of shape constancy, under this view, is intimately connected to the more general problem of shape similarity.

We begin by providing evidence that the perceived constancy of a shape can deteriorate even when one shape is a perspective view of the other, if the amount of perspective distortion becomes large. First, we provide one reasonable definition of perspective distortion in terms of the line at infinity. The line at infinity in the scene plane is defined as the set of points at infinity in any direction. Parallel scene lines may be thought of as intersecting at the line at infinity. Under perspective projection, the line at infinity projects to a real line in the image plane, called the *vanishing line*. In fact, the vanishing line is the intersection of the image plane and a plane parallel to the scene that contains the focal point (see a standard geometry text, such as Coxeter [4] for a more detailed discussion). The distance between the vanishing line and the projection of the scene shape provides a good measure of the image’s

perspective distortion. It tells us, for example, how close to the shape the intersection point will be between image lines that are parallel in the scene. Also, since an image has no absolute scale, we must judge these distances relative to the size of the objects in the image. Therefore, without being very formal or precise in this note, we will define the amount of perspective distortion as the relative distance between the shape in the image and the vanishing line.

Figure 1 shows a planar shape and a perspective view of that shape with the vanishing line relatively close to the shape in the image. It is readily apparent that the shape appears fairly distorted. There is no shape constancy, even though the second image is a true perspective image of the first. We feel that the most likely implication of this is that shape constancy is not always obtained between images related by perspective; rather, shape constancy will be gradually lost as the amount of perspective distortion increases. Pizlo’s experiments [6] do not directly address this possibility.

So far we have considered situations consistent with the theory of Pizlo *et al.* in which one image is a perspective view of another image that has been viewed with known orientation. An alternative plausible definition of shape constancy can be that shape constancy occurs when both images might be perspective views of the same, unknown third planar shape taken with the same camera. This definition has the advantage of treating the images symmetrically, unlike Pizlo *et al.*’s theory, which assumes that the position in space of the object in the first image is known, while the object may appear in any unknown position in the second image. It is shown in [2] that two images are perspective images of the same planar shape if and only if they are related by a projective transformation. This assumes that the images are taken with the same camera. This is a variation on the well-known fact that the set of

* E-mail: ronen@wisdom.weizmann.ac.il. Ronen Basri is an incumbent of the Arye Dissentshik Career Development Chair at the Weizmann Institute.

† E-mail: dwj@research.nj.nec.com.

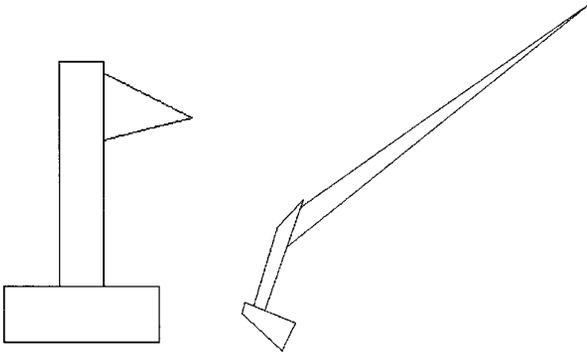


FIG. 1. On the left, an image of a pair slant and tilt. On the right, a perspective view of the same object.

projective views is equivalent to the set of views formed by a series of perspective views taken with cameras that need not be the same. Therefore, if we attempt to remove the asymmetry from Pizlo *et al.*'s theory of shape constancy by relaxing the assumption that viewpoint is known for one image, we find that this leads to the theory that shape constancy is preserved under all projective transformations. However, Pizlo *et al.* and Pizlo [6] have shown persuasively that shape constancy is not preserved under a projective transformation.

These examples suggest that shape constancy is not characterized by a simple class of mathematical transformations. Both examples seem to indicate that we do not judge constancy based on whether two images might be pictures of the same object viewed from any possible viewpoints. Rather, we argue that in judging whether images appear to be of the same object, we rule out extreme viewpoints.

Another question that arises for Pizlo *et al.*'s theory is whether shape constancy can be modeled for 3-D objects. Perspective transformations, which are claimed to be the underlying transformations for shape constancy, properly describe the appearance of planar objects from various viewpoints. Most objects observed by the visual system, however, are nonplanar, and their appearances cannot be accounted for properly by 2-D perspective transformations. Modifications of the theory are therefore necessary if we believe that shape constancy can be obtained for 3-D objects. It is, on the one hand, unclear that one can model shape constancy for 3-D objects using a simple geometric transformation, or on the other hand, that the notion of shape constancy should be restricted to planar shapes.

Human perception of whether two shapes seem the same can be influenced by factors other than the geometric transformation that relate the two shapes. For example, the complexity of a shape may be relevant in constancy judgments; the more complex a shape is, the more tolerant we are of distorting transformations. Figure 2 shows two triangles related by an affine and consequently also by

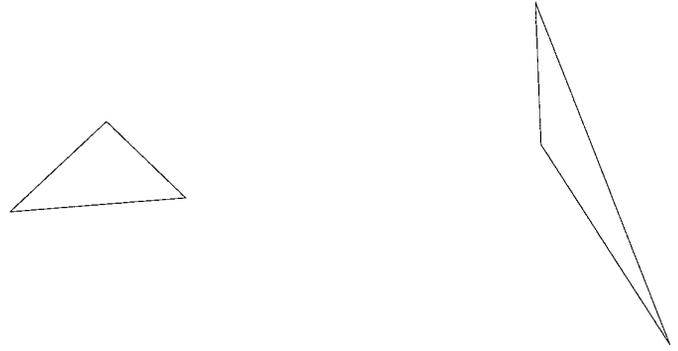


FIG. 2. Two triangles related by an affine transformation.

a perspective transformation (of course all triangles are related by such transformations). Figure 3 shows a more complex shape distorted by the same affine transformation. We suggest that most human observers can perceive the complex shape as the same, related by an affine stretching, while the same observers perceive the triangles as different shapes. This suggests that the class of transformations that we allow, when comparing two shapes for constancy, may depend on the complexity of the shapes.

This may relate to the nonaccidentalness of the relation between the two shapes. Nonaccidental image properties are widely discussed and used in computational vision [3, 5, 7]. An event in an image is considered nonaccidental if the probability that such an event might occur in a random environment is low. In a random world the probability that two arbitrary triangles would be related by an affine transformation is 1, whereas the probability that more complex shapes would be related by an affine transformation is zero. Thus, the two triangles can equally be either two views of the same triangle or simply different triangles. Whereas the two complicated shapes are very unlikely to be two different shapes that happen to be related by an affine transformation. This potentially explains why com-

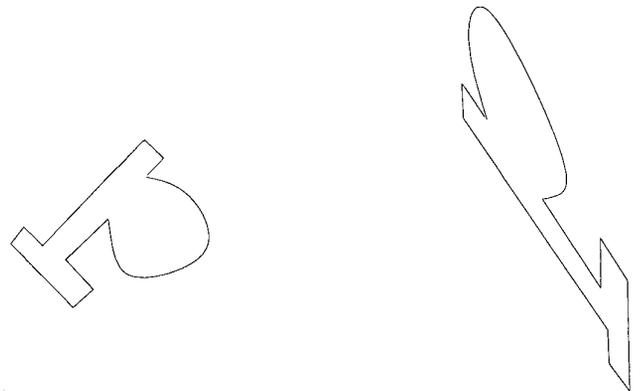


FIG. 3. Two more complex shapes related by the same affine transformation.

plex shapes appear the same when related by an affine transformation.

Context and prior knowledge can also play a role in human perceptions of shape identity. For example, a “p” and “d” can appear to be different, even when they are identical except for a 180° rotation. Perhaps this is because they match different prior models we have of the letters. It seems possible also that context may affect our judgments, by influencing the extent to which we make use of prior knowledge. When we are reading, we do not even notice the similarity of p and d. When we are trying to pick out letters in alphabet soup, to spell a word containing d, we may recognize a p as an upside down d. The context, or the task that we are performing, may also influence the sort of transformations we allow for when comparing two shapes for identity.

In our view, therefore, shape constancy is a complex phenomenon that reflects the interplay of many factors, including the type and magnitude of the viewing transformation, the complexity of the shapes viewed, and context and prior knowledge. This makes shape constancy a difficult problem, but it also means that an understanding of shape constancy can provide insight into a wide range of vision problems. We will conclude this commentary by briefly mentioning the relation of shape constancy to a very general and important problem in human and machine vision, the problem of shape similarity.

Sometimes, two shapes appear to be pictures of the same object, even though they do not appear identical. For example, Fig. 4 appears to show two pictures of a turtle with its limbs bent in different positions. In general, if one distorts a shape a little bit, it appears to be a distorted version of the same shape. After some amount of larger distortion, the shapes no longer appear to be the same. The effect that different types of distortion will have on the similarity of shapes is not well understood. This is clearly, though, a very important problem in vision, underlying the process of recognition, classification, and tracking of nonrigid objects.

It seems likely that the effect of distortions on shape

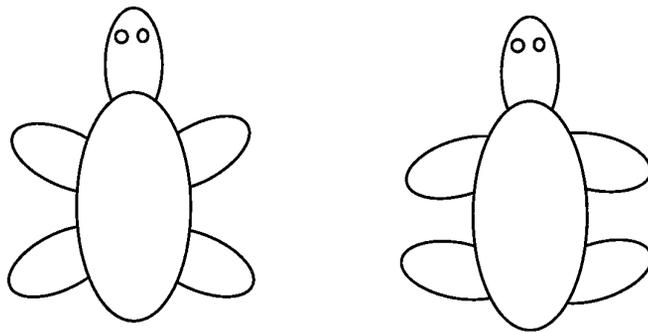


FIG. 4. Two turtles with their limbs in different positions.

similarity depends on many factors, including the amount and type of distortion, the complexity of the object, and context and prior knowledge. We have suggested that these factors also play an important role in the problem of shape constancy. If this is so, it is likely that these problems should be addressed together, and that a complete understanding of the shape constancy problem may require renewed efforts at addressing the problem of shape similarity (see [1] for our efforts at addressing this problem).

REFERENCES

1. R. Basri, J. L. Costa, D. Geiger, and D. W. Jacobs, Determining the similarity of deformable objects, in *IEEE Workshop on Physics-Based Modeling in Computer Vision 1995*, pp. 135–143.
2. R. Basri and D. Jacobs, Constancy and similarity, The Weizmann Institute of Science, technical report, forthcoming.
3. T. Binford, Inferring surfaces from images, *Artif. Intell.* **17**, 1981, 205–244.
4. H. S. H. Coxeter, *The Real Projective Plane*, Springer-Verlag, Berlin/New York, 1993.
5. D. Lowe, *Perceptual Organization and Visual Recognition*, Kluwer Academic, The Netherlands, 1985.
6. Z. Pizlo, A theory of shape constancy based on perspective invariants, *Vision Res.* **34**(12), 1994, 1637–1658.
7. A. Witkin, and J. Tenenbaum, On the role of structure in vision, in *Human and Machine Vision* (Beck, Hope, and Rosenfeld, Eds.), Academic Press, New York, 1983.