VISUAL COGNITION, 2005, 12 (4), 639–690

Psychology Press
Taylor & Francis Group

Commentaries

Stability and change

James Todd
Department of Psychology, Ohio State University, Columbus, USA

Most people would agree that the shape of an object is one of its most perceptually important attributes, and some researchers have argued that it is the primary attribute by which observers are able to recognize objects (e.g., Biederman, 1987). Given the ubiquity of this common intuition, it is somewhat puzzling to note that the concept of ''shape'' has no formal mathematical definition that can adequately characterize its intended meaning when used colloquially. For example, almost everyone would concur that a big sphere and a small sphere both have the same shape, yet by most of the standard measures used in geometry they are quite different.

The abstract nature of the concept of shape is perhaps best revealed by the perceptual classification of biological forms (e.g., see Thompson, 1942). Consider, for example, the ability of normal individuals to identify their friends and loved ones under a variety of different viewing conditions. We are able to identify people from different vantage points, and with different facial expressions, hairstyles, make-up, or clothing accessories, such as hats or jewellery. We can also identify people after they undergo a growth spurt, gain or lose weight, or suffer the effects of ageing. These observations suggest that the identity of an individual's face must be based on some remarkably abstract property that is somehow unaffected by all of the transformations that faces typically undergo in the natural environment.

This core idea of invariance under change has been of great importance to the development of modern geometry. In 1872, the German mathematician Felix Klein gave a lecture at Erlangen University, in which he outlined a general principle for constructing different geometries that is now known as the Erlanger Program. His basic idea was to consider arbitrary groups of single valued transformations, and to investigate the properties of objects they leave invariant. Klein noted that some object properties are more stable than others, because they remain invariant over a larger set of possible transformations. By using different properties to define ''shape equivalence'', it is possible to create a hierarchy of

http://www.tandf.co.uk/journals/pp/13506285.html DOI:10.1080/13506280444000265

geometries (i.e., Euclidean, affine, conformal, etc.) in which structural properties can be stratified with respect to their stability in a formally precise way.

Since the 1980s, Lin Chen has been developing a provocative hypothesis that a similar stratification of object properties may also exist in human perception, and he has accumulated a large body of evidence to support this view, which is elegantly summarized in the preceding target paper. His primary conclusion based on this evidence is that the ordering with which various properties are processed within the human visual system is based on their relative stability. Thus, topological properties are the quickest to be processed, whereas metrical properties are the slowest—precisely the opposite of what would be expected by many current researchers in the field.

I have been sympathetic to Chen's general approach since I first read his report in *Science* on this topic in 1982. At that time, my own research was focused on the perceptual identification of different categories of events. For example, human observers can easily distinguish between rigid and nonrigid motion (Todd, 1982). They can identify a wide variety of gaits in humans and other animals, such as walking, running, skipping, or dancing (Johansson, 1975). They can identify slow viscal-elastic changes in form, such as growth or weight gain (Todd, Mark, Shaw, & Pittenger, 1980), and they are also quite good at identifying intentional, social interactions, such as affection or aggression (Heider & Simmel, 1944). A remarkable property of event perception is that these different types of change can easily be identified even when they are applied to an unexpected structural configuration such as a random pattern of dots. The reason Chen's paper resonated with me in 1982 is that it was closely related to a hypothesis proposed by several other researchers concerning possible sources of information by which different transformations could potentially be identified (see Pittenger & Shaw, 1975; Todd et al., 1980). According to this approach, each type of event may be distinguished by the specific set of object properties it alters, and the specific set of properties it leaves invariant.

A fundamental fact of the natural environment is that all objects undergo change. Human observers possess the ability to identify a specific type of transformation (e.g., rotation) over a wide range of possible objects, and they are also able to identify a specific object (e.g., my wife) over a wide range of possible transformations. Within that context, it should not be surprising that early visual processing should focus primarily on those relatively stable properties that define the essence of an object's identity, rather than those that have a much higher variance under natural viewing conditions. There are many common events in our day-to-day experiences that highlight the importance of stability over change in object recognition. For example, I often fail to notice—at some cost—when my wife comes home with a new hairstyle, but I have never once mistaken her for someone else.

There is another important reason why the human visual system might allocate its resources disproportionately to attributes of objects that have a relatively high degree of stability. Although the objects encountered in natural vision are most often three-dimensional, the visual information by which they are specified is confined to the two-dimensional projection surface of the retina. This projective mapping produces an inevitable loss of information, such that any given pattern of optical stimulation may have an infinity of possible 3-D interpretations. Recent research has shown, however, that these ambiguities can be highly constrained in the sense that all of the possible interpretations are related by a limited class of transformations (see Koenderink, van Doorn, Kappers, & Todd, 2001). Attributes of objects that are invariant over those transformations can be determined quite easily from the available visual information, whereas other properties can only be estimated by incorporating prior knowledge or adopting some ad hoc heuristic. It makes sense, therefore, that the human visual system would concentrate its resources on object attributes that can be determined most reliably.

Although I am in general agreement Chen's overall argument, there is one caveat I would offer about accepting alternative geometries developed by mathematicians too literally as potential models of visual information processing. It is important to keep in mind that mathematics is an idealistic endeavour that demands perfect rigor and consistency. Biology, in contrast, is much more pragmatic. Simple heuristics can evolve quite readily if they impart some benefit to the survival of a species, even if they do not conform to the rigorous standards formal geometry. In Part V of his target paper, Chen presents a well-known figure from Minsky and Papert (1969) that was originally devised to show the computational difficulty of topological relations (see also Todd, Chen, & Norman, 1998). Chen argues that this apparent counterexample to his thesis is due to other psychological factors, rather than a general inability to perceive topological relations. I agree with this assessment. However, by imposing unspecified boundary conditions on the visual processing of topological structure, there is a danger that the ''topology'' of visual perception may turn out to be substantially different from the ''topology'' of formal mathematics.

REFERENCES

Biederman, I. (1987). Recognition-by-components: A theory of human image interpretation. *Psychological Review*, *94*, 115–147.

Chen, L. (1982). Topological structure in visual perception. *Science*, *218*, 699–700.

Chen, L. (this issue). The topological approach to perceptual organization. *Visual Cognition*, *12*, 553–637.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, *57*, 243–259.

Johansson, G. (1975). Visual motion perception. *Scientific American*, *232*, 76–88.

Koenderink, J. J., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2001). Ambiguity and the ''Mental Eye'' in pictorial relief. *Perception*, *30*, 431–448.

Minsky, M. L., & Papert, S. (1969). *Perceptrons: An introduction to computational geometry. Cambridge*, MA: MIT Press.

Pittenger, J. B., & Shaw, R. E. (1975). Aging faces as viscal elastic events: Implications for a theory of nonrigid shape perception. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 374–382.

Thompson, D. W. (1942). *On growth and form*. Cambridge, UK: Cambridge University Press.

Todd, J. T. (1982). Visual information about rigid and nonrigid motion: A geometric analysis. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 238–251.

Todd, J. T., Chen, L., & Norman, J. F. (1998). On the relative salience of Euclidean, affine and topological structure for 3D form discrimination. *Perception*, *27*, 273–282.

Todd, J. T., Mark, L. S., Shaw, R. E., & Pittenger, J. B. (1980). The perception of human growth. *Scientific American*, *242*, 132–144.

Please address all correspondence to: James Todd, Department of Psychology, Ohio State University, 142 Townshend Hall, Columbus, OH 43210, USA. Email: Tod.44@osu.edu

A neural basis for global object features

Robert Desimone

NIH, Bethesda, MD, USA

In my commentary, I will focus on the relationship between Lin Chen's topological model for visual processing and the known biological properties of the visual system.

It is well-established that object recognition in primates is mediated by processing in a hierarchy of visual areas that form the occipitotemporal processing pathway, or ventral stream (see Ungerleider & Mishkin, 1982). Visual processing at the earliest stage (V1) is based on ''local'' (within small receptive fields) analysis of relatively simple visual features, and processing becomes increasingly more ''global'' (large receptive fields), more complex, and more attuned to global object properties as one proceeds along the pathway into the interior temporal (IT) cortex. Most psychological accounts of perception are consistent with the neurophysiological view. Typical psychological models specify that the beginning, or elementary, features of visual perception can be largely identified with the local features processed by neurons in V1. These local features are then combined across different feature domains in extrastriate cortex and then elaborated into representations of figures versus ground and, ultimately, into global representations of complex objects such as faces and houses. How can the neurophysiology and anatomy of perception then be reconciled