Perception, 2009, volume 38, pages 630 – 632

doi:10.1068/p3804rvw

Reviews

3D shape: Its unique place in visual perception by Z Pizlo; MIT Press, Cambridge, MA, 2008, 312 pages, $38.00 (£28.95) ISBN: 978 0 262 16251 7

This book is a very well written and extremely scholarly piece of work. Pizlo's treatment of this highly complex and mathematical subject is both enlightening and easy to read. For the most part he uses skilful illustrations and examples to explain the mathematical concepts involved, relegating the mathematics itself to an appendix. The in-depth, historical nature of the book and the detailed scrutiny applied to past works are truly impressive. There are, however, two things that potential readers should know about this book. First, it is about the processes by which humans apprehend the overall 3D shape of functionally important objects. It is not about how we perceive surface orientation or the detailed undulations within surfaces. Second, it is not a textbook-style summary of received wisdom in shape perception. It is more of a thesis, critically analysing, and in many cases setting aside, past work so as to establish a new theory.

The introductory chapter deals with early theories and experiments; however, it also sets out the key refrains of the book. The first few pages discuss unique importance of shape and its fundamentally complex but structured nature. Pizlo then argues that early attempts to study shape constancy (for example, Thouless 1931a, 1931b) were flawed because the 2D shapes used (ellipses and triangles) were too simple. The projection of an image onto the retina leads to a loss of information, and a given elliptical image can result from any physical ellipse presented with the correct tilt and slant. Thus, Thouless (1931a, 1931b) actually studied shape ambiguity not shape constancy. Correcting this error revealed that the perception of slant and object shape are decoupled (Stavrianos 1945). Thus, Pizlo argues, the often-cited conclusion that humans take slant into account to achieve shape constancy is not valid. The chapter concludes with an extended discussion of the contribution of Gestalt psychology to shape perception. The key contributions are given as the Gestalt emphasis on figure – ground organisation (Wertheimer 1923/1958), the introduction of the simplicity principle, and the discovery that humans can correctly perceive 3D objects from line drawings despite conflicting stereoscopic cues (Kopfermann 1930; Schriever 1925). Thus, chapter 1 introduces the following recurrent themes: (i) the study of shape perception requires the study of shape constancy, (ii) this is best achieved with relatively complex 3D stimuli or 2D representations of 3D objects, (iii) slant has nothing to do with shape constancy, (iv) human shape perception requires figure – ground organisation, and (v) it is facilitated by the application of some form of simplicity principle.

Chapter 2 introduces developments that arose during the Cognitive Revolution. Pizlo divides research conducted in the period into neo-Gestaltism and neo-Empiricism: the former emphasises the role of innate a priori constraints while the latter emphasises the role of experience and learning. Pizlo argues that the neo-Gestaltists made the bigger contribution to our understanding of shape perception. In particular, this group provided a quantitative definition of Gestalt simplicity based on information theory (Shannon 1948; Hochberg and McAlister 1953; Hochberg and Brooks 1960). The Hochberg group went on to show that the ability to perceive 3D objects from 2D pictures is innate. Later, Attneave and Frost (1969) showed that the simplicity of an object and its retinal shape determine the Euclidean properties (specifically edge orientations) of the 3D percept. That is, the perception of edge orientation follows from the perception of shape rather than being a pre-requisite of it.

Chapter 3 discusses the contributions of the machine vision community to the study of shape perception. Machine vision methods for shape recognition introduced during the 1960s and 1970s culminated in the use of model-based invariants and optimisation-based shape reconstruction methods. Both these approaches involve the construction of 3D percepts from 2D shapes on the 'retina' and both require figure – ground organisation to provide the 2D shapes in the first place. Thus figure – ground organisation was recognised as a key problem in machine vision but one that turned out to be very difficult to solve.

Marr's (1982) 2.5D sketch was an attempt to avoid the problem of figure–ground organisation. He proposed, instead, that the perception of 3D shape depends on the prior perception of surface orientations. That is, shape perception involves 'taking slant into account'. Pizlo's conclusion is that Marr's 2.5D sketch was flawed. This view is consistent with the book's recurrent theme that humans do not take surface orientation into account in the perception of object shape. There were alternative views contemporary with the 2.5D sketch; perhaps most notable is that of Barrow and Tenenbaum (1981) who emphasised the role of simplicity-like constraints. This method requires the prior operation of figure–ground organisation and is cited as a forerunner to Pizlo's own model of shape perception.

Chapter 3 also introduces two key contributions of the machine vision community to the study of human vision. The first concerns the role of invariants. Here Pizlo is keen to draw a distinction between mathematical invariants and the perceptual invariants discussed frequently in the work of Gibson (1979). Mathematical invariants describe features that are not changed (ie not lost) by the transformation of the 3D world onto the 2D retina—can such features be used in the reconstruction process? Pizlo argues that perspective (but not projective) invariants can (Pizlo and Rosenfeld 1992). The second contribution concerns the realisation that vision is an inverse problem. The mapping of the 3D world to the 2D retina is a many-to-one mapping. Undoing this transformation in order to work out which of the many possible scenes produced the 2D image requires the application of a priori constraints.

Chapter 4 discusses relatively recent (last 25 years or so) studies of human shape perception, demonstrating how the formalisms of the machine vision community influenced research in human vision. Pizlo notes that efforts in this period were greatly influenced by the work of Marr (1982) and Gibson (1979), who agreed on two points: the importance of surfaces, and the irrelevance of figure–ground organisation. Pizlo thus goes on to describe a body of work aimed at answering the question "can information about the orientation of surfaces be used to perceive 3D shapes?" (page 116). His conclusion is that it cannot. Further, Pizlo presents a body of evidence to suggest that the depth cues (including stereopsis) that are know to contribute to the perception of *surface* shape and orientation make very little if any contribution to the perception of 3D *object* shape.

In chapter 5 Pizlo finally presents his new paradigm for studying shape perception. At the same time he presents an outline of his computational model. Surprisingly little detail of the model is given in the book, but a full description is provided by Li et al (in press). The method can be summarised—very crudely—as involving two key steps. Figure–ground organisation is used to 'identify' 2D shapes on the retina and these shapes are then 'inflated' into 3D objects by applying a number of a priori constraints. These constraints could be regarded as a new formalism of the Gestalt simplicity principle. The second step of the model has been instantiated and it is good at identifying the shape of 3D objects that are broadly symmetrical. Note that it does not deal well with irregular objects such as rocks, but Pizlo would contend that such objects—though 3D—do not possess functional shape. Perhaps more significantly, the first part of the new method, figure–ground organisation, seems to be much less well developed than the latter 3D inflation stage. Although a potential method based on symmetry is discussed (Nevatia 2000), Pizlo can find no method yet to match human vision. Thus, figure–ground organisation remains a core problem in the study of shape perception.

The book ends with a catalogue of the 'milestones' and 'millstones' (Pizlo's terminology) in the study of 3D shape perception. Some may find this categorisation and the somewhat pejorative terms used troublesome. It should be remembered here that Pizlo is talking about *object* shape. It may well be the case that some of his millstones represent milestones in the study of *surface* shape or indeed other aspects of vision. Marr's (1982) 2.5D sketch is criticised, but his desire to test computational theories with psychophysical experiments is undoubtedly honoured in Pizlo's own work.

This book will certainly be of interest to those studying perception of 3D object shape—many may find it challenging. Both the machine and human vision communities can gain from reading this book. The book should be also useful to those studying the perception of surface shape—certainly it will help them to understand what they are not studying. Pizlo includes two interesting asides; one discusses the nature of scientific theory and how the definitions of terms like theory, model, and explanation have changed over the last 150 years. The other explains the separation between machine and human vision in terms of what the two disciplines are trying to achieve, how they go about it, and how they judge success. Doctoral students considering all aspects of vision would do well to take note of these asides.

Andrew Schofield
School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK;
e-mail: A.J.Schofield@bham.ac.uk

References

Attneave F, Frost R, 1969 "Determination of perceived tridimensional orientation by minimum criteria" *Perception & Psychophysics* **6** 391 – 396

Barrow H G, Tenenbaum J M, 1981 "Interpreting line drawings as 3-dimensional surfaces" *Artificial Intelligence* **17** 75 – 116

Gibson J J, 1979 *The Ecological Approach to Visual Perception* (Boston, MA: Houghton Mifflin)

Hochberg J, Brooks V, 1960 "The psychophysics of form—reversible-perspective drawings of spatial objects" *American Journal of Psychology* **73** 337 – 354

Hochberg J, McAlister E, 1953 "A quantitative approach to figural goodness" *Journal of Experimental Psychology* **46** 361 – 364

Kopfermann H, 1930 "Psychologische Untersuchungen über die Wirkung zweidimensionaler Darstellungen körperlicher Gebilde" *Psychologische Forschung* **13** 293 – 364

Li Y, Pizlo Z, Steinman R M, in press, "A computational model that recovers the 3D shape of an object from a single 2D retinal representation" *Vision Research*

Marr D, 1982 *Vision* (New York: W H Freeman)

Nevatia R, 2000 "Perceptual organization for generic object description", in *Perceptual Organization for Artificial Vision Systems* Eds K L Boyer, S Sarkar (Boston, MA: Kluwer) pp 173 – 189

Pizlo Z, Rosenfeld A, 1992 "Recognition of planar shapes from perspective images using contour-based invariants" *Computer Vision, Graphics, and Image Processing: Image Understanding* **56** 330 – 350

Schriever W, 1925 "Experimentelle Studien über stereoskopisches Sehen" *Zeitschrift für Psychologie* **96** 113 – 170

Shannon C E, 1948 "A mathematical theory of communication" *Bell System Technical Journal* **27** 379 – 423

Stavrianos B K, 1945 "The relation of shape perception to explicit judgments of inclination" *Archives of Psychology* **296** 1 – 94

Thouless R H, 1931a "Phenomenal regression to the real object. I" *British Journal of Psychology* **21** 339 – 359

Thouless R H, 1931b "Phenomenal regression to the real object. II" *British Journal of Psychology* **22** 1 – 30

Wertheimer M, 1923/1958 "Principles of perceptual organization", in *Readings in Perception* Eds D C Beardslee, M Wertheimer (New York: van Nostrand) pp 115 – 135

All books for review should be sent to the publishers marked for the attention of the reviews editor. Inclusion in the list of books received does not preclude a full review.