Reviews

Seeing spatial form by M R M Jenkin, L R Harris (Eds); Oxford University Press, Oxford, 2005, 468 pages, £37.99 cloth (US $69.95) ISBN 978 019 5172881

There is no royal road to vision science
This book was published to honour David M Regan, who has made many very important contributions to the psychology and physiology of vision. Specifically, he has contributed to the study of binocular and monocular mechanisms of motion perception, form from luminance, colour, motion, texture, dynamics in colour vision, visually evoked potentials, understanding the magnetic fields of the brain, ophthalmology, vision during driving, aviation and sport, eye movements, and visual development, as well as to single-unit electrophysiology. Prof. Regan has published three single-authored books, edited several, and has published over 200 papers in refereed journals. He has been exceptionally creative, as well as productive. The contents of this book reflect his diversity, creativity, and scientific sophistication very well. This book contains 19 chapters, written by thirty-seven leading investigators, many specialised in the study of spatial form perception. In the present review, we discuss only selected chapters based on our own expertise.

1. Form vision
The second chapter, by Koenderink, van Doorn, and Kappers is titled “Pictorial relief”. The authors start by pointing out that viewing a 2-D picture of a 3-D shape may lead to (i) the percept of a 2-D piece of paper, or (ii) the percept of a 3-D shape. The 3-D percept is not uniquely specified by the 2-D picture, which raises the question whether the 3-D percept is unique, stable, and veridical. By veridical, we mean that the 3-D percept agrees with the 3-D shape that produced the 2-D picture. Koenderink et al concentrated on studying the stability of the 3-D percept — actually its instability — in the presence of changes in its viewing conditions. They did this by measuring the 3-D shape percept with an elliptical probe technique they introduced in an earlier paper (Koenderink et al 1992). The subject’s task was to adjust the aspect ratio and the 2-D orientation of an ellipse so that it was perceived as lying on the perceived 3-D surface. These two parameters uniquely specify the local 3-D surface orientation, so it was natural to assume that collecting a set of the subject’s adjustments would be sufficient to reconstruct the 3-D percept of the surface. The authors characterised the relation between the perceived 3-D surfaces and the conditions under which they were being viewed. They found that when the viewing conditions changed, the perceived 3-D surfaces also changed. These changes could be described by a 3-D affine transformation with four independent parameters: \( x' = x, \ y' = y, \ z' = ax + by + cz + d \) (only the first three parameters, \( a, \ b, \) and \( c \) are of interest here; \( d \) only affects the absolute depth of a 3-D surface, not its shape). Note that this is the most general affine transformation of a 3-D object that does not change the 2-D orthographic image of the 3-D object (the \( x \) and \( y \) coordinates do not change). Why would the visual system restrict the family of possible percepts to 3-D affine transformations? It turns out that this family can be derived theoretically by assuming that the visual system imposes a planarity constraint on the family of 3-D shapes consistent with a 2-D orthographic image of the shapes (Sugihara 1986; Pizlo 2008). Interestingly, the planarity constraint also operates in the case of 3-D shape perception. The difference is that the 3-D shape percept makes use of additional simplicity constraints such as symmetry and maximal 3-D compactness (Li et al 2008; Pizlo 2008).

Chapter 3, by Westheimer, continues a discussion of the value of using geometrical models in visual perception that has been under way for a very long time. Since the 19th century, mathematicians have formulated a number of geometries that go far beyond ‘classical’ Euclidean geometry. They differ considerably in both their generality and properties. It is useful to note that some developments in geometry, which were formulated without any explicit motivation from physics, were eventually adopted by physicists to explain a number of natural phenomena. Perhaps the best-known example is Einstein’s use of a non-Euclidean geometry to model the curved space in his general relativity theory. Could something similar be successful for developing theories
of visual space perception? Westheimer hopes that it could, and chooses Luneburg’s (1947) theory
of binocular space as an example of how this kind of approach could be useful in visual perception.
Luneburg, using data from experiments on binocular settings of equidistant and parallel lines,
showed that a non-Euclidean space with a constant negative curvature could account well for the
systematic errors of his observers. At the time of its conception, the importance of Luneburg’s
theory might have primarily rested on the fact that it used the same mathematical tools that were
being used by physicists. Today, this theory may seem less important to many vision scientists
because it can handle only a very limited range of data obtained under very unnatural condi-
tions. Let us hope that Westheimer’s call for renewed efforts towards formulating geometrical
models of visual space will inspire new and exciting developments along the lines he proposes in
his chapter.

The last chapter in this section is by Tyler. His chapter begins by pointing out that the ultimate
goal of visual perception is the perception of 3-D objects. According to Tyler, the structure (form
or shape) of a 3-D object is represented by contours that are interpolated by surfaces. We perceive
surfaces even when the surfaces have no features. According to Tyler, surfaces are critical in
determining the percept of an object and its shape because they represent the boundary between
the inside and outside of the object. Surfaces actually tell the observer where the object is. If
surfaces are important characteristics of the percept even when they are not explicitly represented
in the 2-D retinal image, two questions arise: How are surfaces represented in the visual system?
How are they derived from the incomplete information available in the 2-D retinal image?
Surfaces must be inferred (guessed) from other sources of information, such as contours, because
surfaces are not always represented explicitly in the 2-D retinal image, and, when they are
represented explicitly by the presence of texture, texture is not sufficient for the 3-D shape to be
reconstructed reliably (eg Pizlo and Salach-Golyska 1995). Technically, 3-D surfaces result from
solving an inverse problem by imposing a priori constraints on the family of possible 3-D inter-
pretations (Pizlo 2001). The smoothness of the surface is probably the most important constraint
leading to this solution. Kanizsa’s figures, particularly the 3-D versions developed by Tse (1999),
are nice illustrations of this kind of inferential process. Tyler suggests that visually reconstructed
surfaces are minimal surfaces, subject to the constraints provided by their contours. He illustrates
his suggestion by showing how such a surface can be produced with a Kanizsa triangle as his
stimulus.

2. Motion and colour
Studying visuo-motor coordination in the case of driving, flying, and sport is especially interesting,
because it tests what the visual system can do at its limits. Such studies are described in chap-
ter 9 by Gray. Accurate actions in flying, driving, and baseball have to be executed under very
tight time constraints. Results and observations that come from such studies provide the most
challenging benchmarks for machine-vision and robotics community. There are algorithms that
can read text, solve difficult problems such as the Traveling Salesman Problem and even play
chess, but it will take a long time before a machine can beat a good tennis player, or drive a car
on a busy street. The underlying algorithms must work very well before they may be considered
possible models of human perception and action. As an example of research presented by Gray,
consider how pilots estimate the time-to-contact (TTC) during landing. Gray begins with an
orthographic approximation to the perspective projection (perpendicular approach). In particular,
his formula (9.5) assumes that the retinal image of the runway is rectangular, rather than trape-
zoidal, that the approach velocity is constant, the approach path is a straight line and the visual
angle of the width of the runway is small. An orthographic approximation generalises to a
perspective projection (Gray’s angled approach), but only when the visual angle of the runway
is measured at the point on the runway where the plane will actually land. It is not clear, however,
how well the pilot can actually judge the point of landing. If he/she over- or under-estimates
this point, TTC will be overestimated or underestimated. Is there information on the retina about
the landing point? The answer is ‘yes’. Specifically, the vanishing point of the runway can provide
accurate information about the landing point. If the plane’s trajectory is a straight line which
forms an angle \( \alpha \) with the ground, then the visual angle formed by the vanishing point, the pilot’s
eye, and the point of landing is also \( \alpha \) during the entire approach. If the pilot is able to judge
the position of the vanishing point (ie of the horizon), then he/she might be able to judge the point of
landing, as well. Note that if the observer can use the information about the vanishing point, then perhaps he/she can use information about the entire shape of the retinal image of the runway. In other words, one can try to derive an estimate of TTC from temporal changes in the shape of the retinal (perspective) image of the runway, rather than from temporal changes of the width of the runway (as was done in the orthographic approximation).

Spillmann, Pinna, and Werner (chapter 10) tested the role of the watercolour effect in figure–ground organisation. The watercolour effect refers to the fact that when a colour boundary is placed right inside a contour surrounding a 2-D region in the image, the entire region is perceived as coloured, despite the fact that it is white. This is illustrated in figures 1a and 1b. They showed that the watercolour effect determines the figure–ground organisation, that is, that the area, which is perceived as coloured, is also perceived as the figure. Furthermore, they showed that the watercolour effect is a stronger determinant of figure–ground organisation than any of the conventional Gestalt factors, such as proximity, continuity, and familiarity. The authors then suggest that the watercolour effect, which determines figure–ground organisation, can be explained by the perception of the 3-D convex surface of the area in which the colour is spreading. There may, however, be another factor affecting figure–ground organisation: it has been known for some time that local 2-D convexity is a strong factor in determining which side of a 2-D contour is perceived as figure (Kanizsa and Gerbino 1976; Peterson and Gibson 1994).

Now consider the kinds of stimuli used by Spillmann et al: figures 1a and 1b illustrate the stimuli that produced their watercolour effect. They are like those shown in chapter 10. Note that the image in figure 1b was produced by exchanging the colours of the contours in figure 1a. The blue contour, whose luminance contrast is higher, is perceived as the boundary between the figure and its background. Also, note that the smooth (locally convex) parts of this contour face opposite sides in figures 1a and 1b. This becomes very clear in figures 1c and 1d where only the blue contours are shown. Clearly, local 2-D convexity is a confounding factor in the stimuli representing colour-spreading in these stimuli. Fortunately, this confound, present in the stimuli used by Spillmann et al, does not eliminate colour-spreading as a factor in figure–ground organisation completely because it is known that, when the contours are straight, rather than wavy, line segments, figure–ground organisation is still affected by the watercolour effect (Pinna et al 2001). But, the effect with straight-line contours is substantially weaker than the effects reported by Spillmann et al, a fact which should be kept in mind when evaluating the relative importance of the watercolour effect in figure–ground organisation as described in this chapter.

Figure 1. [In colour online, see http://www.perceptionweb.com/abstract/cgi?id=p3410rvw] (a) and (b) illustrate the watercolour effect, (b) was generated by exchanging colours in (a), and (c) and (d) allow an evaluation of figure–ground organisation when only the blue contours in (a) and (b) are present.
3. Eye movements
Chapters 11 and 12 describe research on several aspects of eye movements. The oculomotor system is special among motor systems in that it is most closely related to visual perception. Whereas one can study a number of aspects of locomotion or hand movements without visual input, eye movements are almost always studied with visual input, the natural, or 'adequate' stimulus for initiating and guiding them. Eye (and head) movements must be used to obtain information from within the visual field because the 'centre of best vision', near the centre of the fovea, is very small. Moving the head and the eyes can be essential whenever finding and seeing fine visual details is important. Visual information is needed to plan and execute many different kinds of physical activities, activities as different as reaching for objects, walking, driving, flying, and playing baseball. They are also required to perform very different kinds of mental operations such as finding a tumor in an X-ray, reading, solving physics or math problems, and contemplating a work of art.

The chapter by Kowler discusses the relation between saccadic eye movements and visual search, especially their role during tasks requiring cognitive processing. Specifically, she examines how the decision is made where to move the line-of-sight and how quickly to move it from the current to the next position within the visual field when it must be moved to acquire information. Her main result is that saccadic eye movements, the high-velocity eye movements used to shift gaze rapidly, are not planned too carefully. She explains her results by pointing out that planning and analysing visual information requires time as well as attentional resources. Time spent planning saccades would interfere with cognitive processing, so subjects often choose to make several relatively inaccurate saccades to get the line-of-sight near to where the target might be, rather than devoting resources to getting the line-of sight where the target is most likely to be. In other words, using visual attention and/or decision-processing for saccadic planning and action is costly, so it is sacrificed in order to facilitate cognitive processing.

In chapter 12, Steinman, Menezes, and Herst describe their studies of eye movements during coordinated visuomotor acts made with a unique head-and-eye-movement recording system, a system that makes accurate binocular recordings of the direction of gaze, with the head and upper body free to move naturally. This chapter offers an unusual combination of text, PowerPoint slides, and demos on the CD accompanying the book. The authors begin by pointing out how important (critical) it is to measure oculomotor performance under natural conditions, that is, when the head is free to move and when recordings are made binocularly. Next, they describe results indicating that the control of gaze is quite different during a visuomotor task that actually serves a useful purpose (tapping targets in a given order) than it is when the subject shifts gaze for its own sake, that is, when the subject only looks at a series of targets in a given order. These experiments show clearly that the oculomotor system, observed when realistic tasks are performed under relatively natural conditions, performs very differently than it does when it is studied monocularly with the head immobilised. The second part of the chapter describes recordings made when the subject performed real-life activities, such as grooming a toy monkey, using utensils to eat, knapping a slate arrowhead and assembling a Barbie Doll. Recordings were also made while subjects did a variety of natural tasks while they were jostled about as if they were riding on a bumpy road. One of the striking results in this recent work was that the 'cyclopean' fixation 'point' was often well beyond the target of interest, particularly when the subject searched for an object needed to perform a task despite knowing that the object was nearby. [The cyclopean fixation point is the point closest (in the least-squares sense) to the lines of sight of the two eyes.] Furthermore, when there were large fixation errors in each of the eyes, as there often were, the direction of the cyclopean gaze was quite accurate!

4. Neural basis of form vision
This section begins with Regan and Regan's chapter. It provides a brief description of their papers in which electrical and magnetic brain activity served as the experimental data. Each contribution is described briefly, but there are so many contributions that this chapter required more than 60 pages! Their story begins when David ('Martin') Regan's career began. He began by obtaining the background in physics that prepared him exceptionally well for research in both physiology and psychophysics, research that has contributed so much to our present understanding of visual perception. He started by designing an extremely sensitive method for recording steady-state evoked
potentials in response to visual stimuli. This research led to his doctoral dissertation in physics. The chapter next describes studies of responses evoked by flickering isoluminant stimuli. This pioneering work subsequently elucidated the properties of the physiological mechanisms underlying colour-defined form. He, and his collaborator, Marian Regan, then describe an experiment in which evoked potentials were recorded in response to depth- and texture-defined form. They next explored the relation between the temporal and spatial aspects of vision. The chapter ends with a description of one of Regan and Regan’s most important contributions: their demonstration of how a nonlinear systems analysis can be used to mathematically identify (understand) the various stages of neural processing. Studying the input–output relations of a linear ‘black box’ does not allow one to figure out the contents of the box, but when the box is a nonlinear system, the input–output relations are much more informative. The brain is highly nonlinear. Therefore, treating it as such, despite the mathematical sophistication required to do this, offers two advantages: (i) an adequate approach, and (ii) much greater explanatory power. Visual science would benefit enormously if the approach described in this chapter were to be adopted universally in studies of early stages of visual processing.

We close our review by saying that this book is stimulating as well as very informative. In this respect it resembles the career of the outstanding scientist being honoured. The chapters cover an enormous range of topics. Individual chapters can easily be used to supplement a graduate course or seminar on visual perception. Reading the book is not easy, in part, because of the diversity of topics, and, in most cases, because of the advanced level on which the material is presented. The reader is assumed to have background in mathematics and the physiology of the brain, as well as in perception. Clearly, the reader will have to put effort into reading and understanding this book, but it is surely worth it. Paraphrasing Euclid, “there is no royal road to vision science”.

Acknowledgments. We were supported by grants from the National Science Foundation and the US Department of Energy.

Tadamasa Sawada, Zygmunt Pizlo
Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907-2081, USA; e-mail: pizlo@psych.purdue.edu

References
Kanizsa G, Gerbino W, 1976 “Convexity and symmetry in figure-ground organization”, in Vision and Artifact Ed. M Henle (New York: Springer) pp 25–32
Koenderink J J, Doorn A J van, Kappers A M L, 1992 “Surface perception in pictures” Perception & Psychophysics 52 487 – 496
Li Y, Pizlo Z, Steinman R M, 2008 “A computational model that recovers the 3D shape of an object from a single 2D retinal representation” Vision Research (in press)
Luneburg R K, 1947 Mathematical Analysis of Binocular Vision (Princeton, NJ: Princeton University Press)
Peterson M A, Gibson B S, 1994, “Must figure–ground organization precede object recognition?” Psychological Science 5 253 – 259
Pinna B, Breïlstaff G, Spillmann L, 2001 “Surface color from boundaries: a new ‘watercolor’ illusion” Vision Research 41 2669 – 2676
Pizlo Z, 2001 “Perception viewed as an inverse problem” Vision Research 41 3145 – 3161
Pizlo Z, 2008 3D Shape: Its Unique Place in Visual Perception (Cambridge, MA: MIT Press)
Pizlo Z, Salach-Golyska M, 1995 “3D shape perception” Perception & Psychophysics 57 692 – 714
Sugihara K, 1986 Machine Interpretation of Line Drawings (Cambridge, MA: MIT Press)
Tse P, 1999 “Volume completion” Cognitive Psychology 39 37 – 68

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.