

The visual perception of 3D shape[☆]

James T. Todd

Department of Psychology, Ohio State University, Columbus, Ohio, 43210, USA

A fundamental problem for the visual perception of 3D shape is that patterns of optical stimulation are inherently ambiguous. Recent mathematical analyses have shown, however, that these ambiguities can be highly constrained, so that many aspects of 3D structure are uniquely specified even though others might be underdetermined. Empirical results with human observers reveal a similar pattern of performance. Judgments about 3D shape are often systematically distorted relative to the actual structure of an observed scene, but these distortions are typically constrained to a limited class of transformations. These findings suggest that the perceptual representation of 3D shape involves a relatively abstract data structure that is based primarily on qualitative properties that can be reliably determined from visual information.

One of the most remarkable phenomena in the study of human vision is the ability of observers to perceive the 3D shapes of objects from patterns of light that project onto the retina. Indeed, were it not for our own perceptual experiences, it would be tempting to conclude that the visual perception of 3D shape is a mathematical impossibility, because the properties of optical stimulation appear to have so little in common with the properties of real objects encountered in nature. Whereas real objects exist in 3-dimensional space and are composed of tangible substances such as earth, metal or flesh, an optical image of an object is confined to a 2-dimensional projection surface and consists of nothing more than patterns of light. Nevertheless, for many animals, including humans, these seemingly uninterpretable patterns of light are the primary source of sensory information about the arrangement of objects and surfaces in the surrounding environment.

Scientists and philosophers have speculated about the nature of 3D shape perception for over two millennia, yet it remains an active area of research involving many different disciplines, including psychology, neuroscience, computer science, physics and mathematics. The present article reviews the current state of the field from the perspective of human vision: it will first summarize how patterns of light at a point of observation are mathematically related to the structure of the physical environment; it will then consider some recent psychophysical findings on the nature of 3D shape perception; it will evaluate some possible data structures by which 3D shapes might be

perceptually represented; and it will summarize recent research on the neural processing of 3D shape.

Sources of information about 3D shape

There are many different aspects of optical stimulation that are known to provide perceptually salient information about 3D shape. Several of these properties are exemplified in Figure 1. They include variations of image intensity or shading, gradients of optical texture from patterns of polka dots or surface contours, and line drawings that depict the edges and vertices of objects. Other sources of visual information are defined by systematic transformations among multiple images, including the disparity between each eye's view in binocular vision, and the optical deformations that occur when objects are observed in motion.

How is it that the human visual system is able to make use of these different types of image structure to obtain perceptual knowledge about 3D shape? The first formal attempts to address this issue were proposed by James Gibson and his students in the 1950s [1]. Gibson argued that to perceive a property of the physical environment,

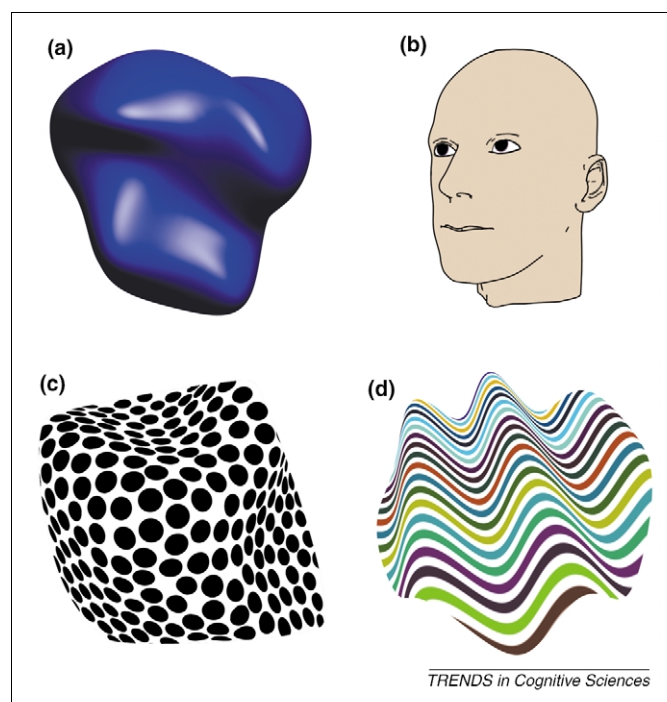


Figure 1. Some possible sources of visual information for the depiction of 3D shape. The 3D shapes in the four different panels are perceptually specified by: (a) a pattern of image shading, (b) a pattern of lines that mark an object's occlusion contours and edges with high curvature, (c) gradients of optical texture from a pattern of random polka dots, and (d) gradients of texture from a pattern of parallel surface contours.

[☆] Supplementary data associated with this article can be found at doi: 10.1016/j.tics.2004.01.006

Corresponding author: James T. Todd (todd.44@osu.edu).

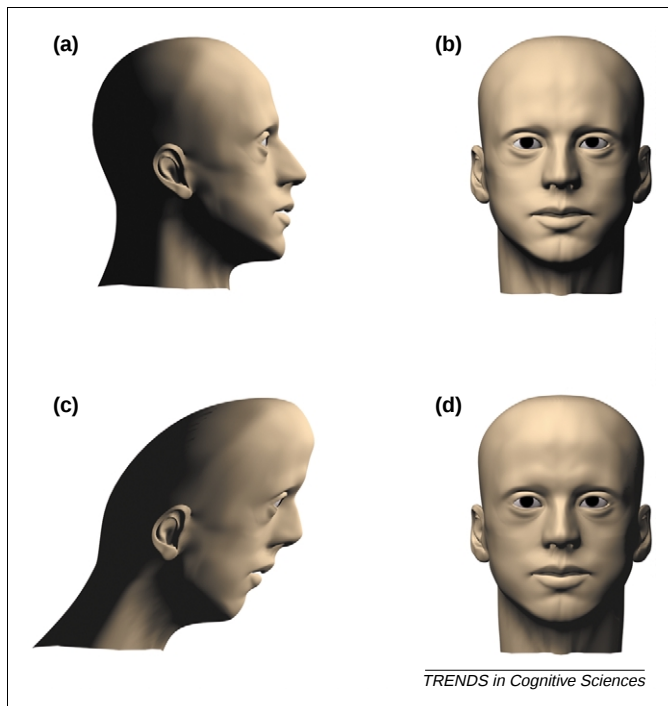


Figure 2. The inherent ambiguity of image shading. Belumeur, Kriegman and Yuille [4] have shown that the pattern of shadows in an image is inherently ambiguous up to a stretching or shearing transformation in depth. (a) and (b) show the front and side views of a normal human head. (c) and (d), by contrast, show front and side views of this same head after it and the light source were subjected to an affine shearing transformation. Note that the two front views are virtually indistinguishable, even though the depicted 3D shapes are quite different.

it must have a one-to-one correspondence with some measurable property of optical stimulation. According to this view, the problem of 3D shape perception is to invert (or partially invert) a function of the following form: $\Lambda = f(\phi)$, where ϕ is the space of environmental properties that can be perceived, and Λ is the space of measurable image properties that provide the relevant optical information. The primary difficulty for this approach is that in most natural contexts the relation between ϕ and Λ is almost always a many-to-one mapping: that is to say, for any give pattern of optical stimulation, there is usually an infinity of possible 3D structures that could potentially have produced it.

The traditional way of dealing with this problem is to assume the existence of environmental constraints to restrict the set of possible interpretations. For example, to analyze the pattern of texture in panel c of Figure 1, it would typically be assumed that the actual surface markings are all circular, and that the shape variations within the 2D image are due entirely to the effects of foreshortening from variations of surface orientation [2]. Similarly, an analysis of the contour texture in panel d of Figure 1 would most likely assume that the actual contours carve up the surface into a series of parallel planar cuts [3]. Although some of the constraints that have been invoked to resolve ambiguities in the visual mapping are intuitively quite plausible, there are others that have been adopted more for their mathematical convenience than for their ecological validity. The problem with this approach is that the resulting analyses of 3D shape might only function effectively within

narrowly defined contexts, which have a small probability of occurrence in the natural environments of real biological organisms.

The inherent ambiguity of visual information is not always as serious a problem as it might appear at first. Although a given pattern of optical stimulation can have an infinity of possible 3D interpretations, it is often the case that those interpretations are highly constrained, such that they are all related by a limited class of transformations. Recent theoretical analyses have shown, for example, that the optical flow of 2-frame motion sequences or the pattern of shadows in an image are ambiguous up to a set of stretching or shearing transformations in depth (see Figure 2) [4,5]. It is important to note that these are linear transformations – sometimes called ‘affine’ – that preserve a variety of structural properties, such as the relative signs of curvature on a surface, the parallelism of lines or planes, and relative distance intervals in parallel directions. Thus, 2-frame motion sequences or patterns of shadows can accurately specify all aspects of 3D shape that are invariant over affine transformations, but they cannot specify other aspects of 3D structure involving metrical relations among distance intervals in different directions.

It is important to point out in this context that there are some potential sources of visual information by which it is theoretically possible to determine the 3D shape of an object unambiguously. These include apparent motion sequences with three or more distinct views [6] and binocular displays with both horizontal and vertical disparities [7]. Because motion and stereo are such powerful sources of information, especially when presented in combination [8], it should not be surprising that they are of primary importance for the perception of 3D shape in natural vision. However, there is a large body of empirical evidence to indicate that human observers are generally incapable of making full use of that information. When required to make judgments about 3D metric structure from moving or stereoscopic displays, observers almost always produce large systematic errors (see [9] for a recent review).

Psychophysical investigations of 3D shape perception

The earliest psychophysical experiments on perceived 3D shape were performed in the 19th century to investigate stereoscopic vision, although the stimuli used were generally restricted to small points of light presented in otherwise total darkness. These studies revealed that observers’ perceptions can be systematically distorted, such that physically straight lines in the environment can appear perceptually to be curved, and apparent intervals in depth become systematically compressed with increased viewing distance. Given the impoverished nature of the available information in these experiments, it is reasonable to be skeptical about the generality of their results, but more recent research has shown that these same patterns of distortion also occur for judgments of real objects in fully illuminated natural environments [9–14]. A particularly compelling example can be experienced while driving in an automobile along a multi-lane highway. In the United States, the hash marks that separate passing lanes all have a length of 10 ft (3.05 m), but that is

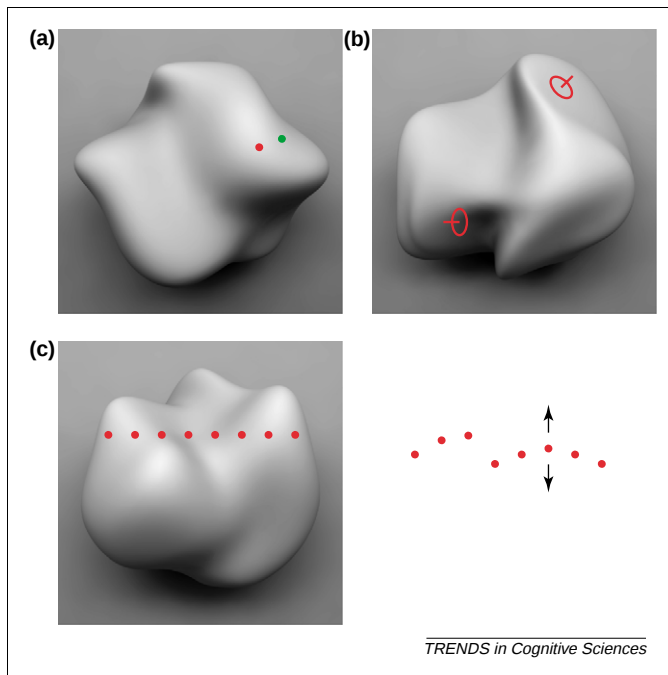


Figure 3. Alternative methods for the psychophysical measurement of perceived 3D shape. (a) depicts a possible stimulus for a relative depth probe task. On each trial, observers must indicate by pressing an appropriate response key whether the red dot or the green dot appears closer in depth. (b) shows a common procedure for making judgments about local surface orientation. On each trial, observers are required to adjust the orientation of a circular disk until it appears to fit within the tangent plane of the depicted surface. Note that the probe on the upper right of the object appears to satisfy this criterion, but that the one on the lower left does not. (c) shows a possible stimulus for a depth-profile adjustment task. On each trial, an image of an object is presented with a linear configuration of equally spaced red dots superimposed on its surface. An identical row of dots is also presented against a blank background on a separate monitor, each of which can be moved in a perpendicular direction with a hand held mouse. Observers are required to adjust the dots on the second monitor to match the apparent surface profile in depth along the designated cross-section. By obtaining multiple judgments at many different locations on the same object, it is possible with all of these procedures to compute a specific surface that is maximally consistent in a least-squares sense with the overall pattern of an observer's judgments [18].

not how they appear perceptually to human observers. Those in the distance appear much shorter than those closer to the observer.

Until quite recently, most experiments on the perception of 3D shape have used relatively crude psychophysical measures, such as judging the magnitude of an object's extension in depth, or estimating the ratio of its depth and width. It is important to recognize that this type of procedure is obviously inadequate for revealing the richness of human perception. After all, a sphere, a cube or a pyramid can have identical depths and widths, yet all observers would agree that their shapes are quite different. During the past decade our empirical understanding about 3D shape perception has been significantly enhanced by the development of more sophisticated psychophysical methods [15–19], several of which are described in Figure 3. What they all share in common is that observers are required to estimate some aspect of local 3D structure at many different probe points on an object's surface, and these responses are then analyzed to compute a surface that is maximally consistent in a least-squares sense with the overall pattern of an observer's judgments.

Consider, for example, a recent experiment by Todd and co-workers [20]. Observers in this study made profile adjustments (Figure 3c) for images of randomly shaped surfaces similar to the one in the lower left panel of Figure 1 that were depicted with different types of texture. An analysis of these judgments revealed that they were almost perfectly correlated with the simulated 3D structures of the depicted surfaces. The correlation between different observers was also high, as was the test-retest reliability for individual observers across multiple experimental sessions. These findings indicate that observers' judgments about the general pattern of concavities and convexities were quite accurate, but there was one aspect of the apparent 3D structure that was systematically distorted. The judged magnitude of relief was underestimated by all observers, and there were large individual differences in the extent of this underestimation that ranged from 38% to 75%. These results suggest that the available information from texture gradients can only specify the 3D shape of a surface up to an indeterminate depth scaling. Thus, when observers are required by experimental task demands to estimate a specific magnitude of relief, they are forced to guess or to adopt some ad hoc heuristic.

Although this general pattern of results is quite common in experiments on 3D shape perception, it is by no means universal [19,21,22]. Sometimes the variations among observers' judgments are more complex and cannot be accounted for by a simple depth scaling transformation. A particularly clear example of this phenomenon has recently been reported by Koenderink *et al.* [18]. The stimuli in their study consisted of four shaded photographs of abstract sculptures by the Romanian artist Constantin Brancusi. Over a series of experimental sessions, the depicted shape in each photograph was judged by observers using all of the different methods described in Figure 3. These judgments were then analyzed to compute a best-fitting response surface for each observer in each condition, and these response surfaces were compared using regression analyses. A particularly surprising result from this study is that in some instances the correlations between the judged 3D structures obtained by a given observer using different response tasks were close to zero! Subsequent analyses revealed, however, that almost all of the variance between these conditions could be accounted for by an affine shearing transformation in depth like the one depicted in Figure 2.

It is important to recognize that this pattern of results is perfectly consistent with the inherent ambiguity of shading information described by Belumeur *et al.* [4]. These findings indicate that when observers make judgments about objects depicted in shaded images, they are quite accurate at estimating those aspects of structure that are uniquely specified by the available information – in other words, those that are invariant over affine stretching or shearing transformations in depth (Figure 2). Because all remaining aspects of structure are inherently ambiguous, observers must subconsciously adopt some type strategy to constrain their responses, and it appears that these strategies do not necessarily remain constant

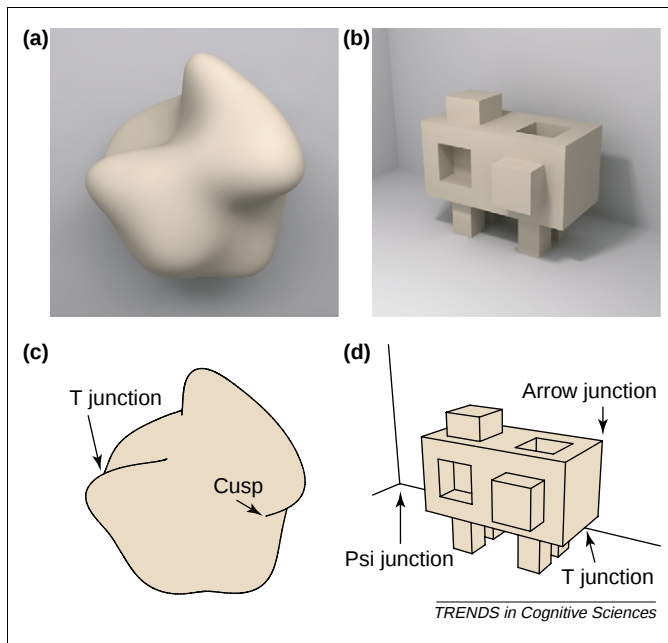


Figure 4. Some important features of local surface structure that can provide perceptually useful information about 3D shape even within schematic line drawings. (a) and (b) show shaded images of a smoothly curved surface and a plane-faced polyhedron. (c) and (d) show schematic line drawings of these scenes in which the lines denote occlusion contours or edges of high curvature. Several different types of singular points (identified with arrows) are particularly informative for specifying the qualitative 3D structure of an observed scene. A more complete analysis of these different types of image features is described in a classic paper by Malik [30].

across different response tasks. A similar result has also been reported by Cornelis *et al.* [21], who found affine shearing distortions between the judged 3D shapes of objects depicted in images viewed at different orientations.

The perceptual representation of 3D shape

Almost all existing theoretical models for computing the 3D structures of arbitrary surfaces from visual information are designed to generate a particular form of data structure that can be referred to generically as a 'local property map'. The basic idea is quite simple and powerful. A visual scene is broken up into a matrix of small local neighborhoods, each of which is characterized by a number (or a set of numbers) to represent some particular local aspect of 3D structure, such as depth or orientation. This idea was first proposed over 50 years ago by James Gibson [23], although he eventually rejected it as one of his biggest mistakes [24].

One major shortcoming of local property maps as a possible data structure for the perceptual representation of 3D shape is that they are highly unstable. Consider what occurs, for example, when an object is viewed from multiple vantage points. In general, when an object moves relative to the observer (or vice versa), the depths and orientations of each visible surface point will change, so that any local property map that is based on those attributes will not exhibit the phenomenon of shape constancy. In principle, one could perform a rigid transformation of the perceived structure at different vantage points to see if they match. This would only work, however, if the perceived metric structure were veridical, and the empirical evidence shows quite clearly that is not the case.

What type of data structure could potentially capture the qualitative aspects of 3D surface shape without also requiring an accurate or stable representation of local metric properties? It has long been recognized that a convincing pictorial representation of an object can sometimes be achieved by drawing just a few salient features (see Figure 4). For example, one such feature that is especially important is the occlusion contour that separates an object from its background [25]. Indeed, an occlusion contour presented in isolation (e.g. a silhouette) can often provide sufficient information to recognize an object, and to reliably segment it into distinct parts [26–28]. Another class of features that is perceptually important for segmentation and recognition includes the edges and vertices of polyhedral surfaces [29,30], and there is some evidence to suggest that the topological arrangement of these features provides a relatively stable data structure that can facilitate the phenomenal experience of shape constancy (see Box 1).

Within the literature on both human and machine vision, there have been numerous attempts to analyze line drawings of 3D scenes. This research was initially focused on the interpretation of line drawings of simple plane-faced polyhedra [31,32]. Researchers were able to exhaustively catalog the different types of vertices that can arise in line drawings of these objects, and then use that catalog to label which lines in a drawing correspond to convex, concave, or occluding edges. Similar procedures were later developed to deal with other types of lines corresponding to shadows or cracks, and the occlusion contours of smoothly curved surfaces [30]. A closely related approach has also been used to segment objects into parts, which can be distinguished from one another by different types of features of which they are composed. The arrangement of these parts provides sufficient information to successfully recognize a wide variety of common 3D objects [33]. Moreover, because the classification of vertices and edges is generally unaffected by small changes in 3D orientation, this method of recognition has a high degree of viewpoint-invariance relative to other approaches that have been proposed in the literature.

There is an abundance of evidence from pictorial art and human psychophysics that occlusion contours and edges of high curvature play an important role in the perception of 3D shape, but the mechanisms by which these features are identified within 2D images remain poorly understood. One important reason why edge labeling is so difficult is that the pattern of 2D image structure can be influenced by a wide variety of environmental factors. Occlusion contours and edges of high curvature are most often specified by abrupt changes of image intensity, but similar abrupt changes can also be produced by cast shadows, specular highlights, or changes in surface reflectance (see Figures 1 and 4). Human observers generally have little difficulty identifying these features, but there are no formal algorithms that are capable of achieving comparable performance, despite over three decades of active research on this problem.

The identification of image features is also of crucial importance for traditional computational analyses of 3D structure from motion or binocular disparity. These analyses are all based on a fundamental assumption that

Box 1. Sources of perceptual constancy

An important topic in the theoretical analysis of visual information is to identify informative properties of optical structure that remain stable over changing viewing directions. For example, researchers have shown that the terminations of edges and occlusion contours in an image have a highly stable topological structure. Although these features can sometimes appear or disappear suddenly, these transitions are highly constrained and only occur in a few possible ways, which have been exhaustively enumerated [57]. Tarr and Kriegman [58] have recently demonstrated that the occurrence of these abrupt events can dramatically improve the ability of observers to detect small changes in object orientation.

Stability over change can also be important for the perceptual representation of 3D shape. Unlike other aspects of local surface structure (e.g. depth or orientation), curvature is an intrinsically defined attribute that does not require an external frame of reference. Thus, because it provides a high degree of viewpoint-invariance, a curvature-based representation of 3D shape could be especially useful for achieving perceptual constancy. Several sources of evidence suggest that local maxima or minima of curvature provide important landmarks in the perceptual organization of 3D surface structure. For example, when observers are asked to segment an object's occlusion boundary into perceptually distinct parts, they most often localize the part boundaries at local extrema of negative curvature [26–28]. A similar pattern of results is obtained when observers are asked to place markers along the ridges and valleys of a surface [59]. These judgments remain remarkably stable over changes in surface orientation, and the marked points generally coincide with local maxima of curvature for ridges and

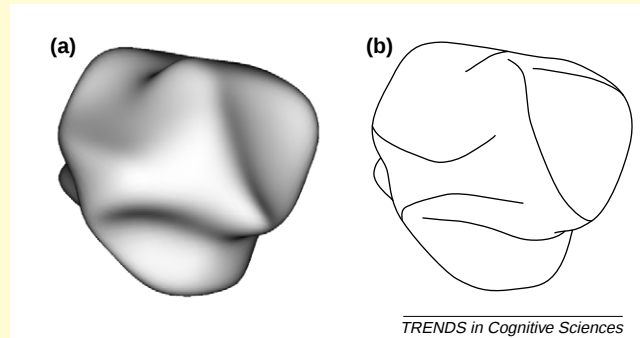


Figure 1. Two methods of pictorial depiction of smoothly curved surfaces. (a) a shaded image of a randomly shaped object; (b) the same object depicted as a line drawing. The lines on the figure denote two different types of surface features: smooth occlusion contours, where the surface orientation at each point is perpendicular to the line of sight, and curvature ridge lines, where the surface curvature perpendicular to the contour is a local maximum or minimum. The configuration of contours provides compelling information about the overall pattern of 3D shape.

local minima of curvature for valleys. There is other anecdotal evidence to suggest, moreover, that the depiction of smooth surfaces in line drawings can be perceptually enhanced by the inclusion of curvature ridge lines (see Figure 1).

visual features must projectively correspond to fixed locations on an object's surface. Although this assumption is satisfied for the motions or binocular disparities of textured surfaces, it is often strongly violated for other types of visual features, such as smooth occlusion boundaries, specular highlights or patterns of smooth shading. There is a growing amount of evidence to suggest, however, that the optical deformations of these features do not pose an impediment to perception, but rather, provide powerful sources of information for the perceptual analysis of 3D shape [34]. Some example videos that demonstrate the perceptual effects of these deformations are provided in the supplementary materials to this article.¹

The neural processing of 3D shape

Although most of our current knowledge about the perception of 3D shape has come from computational analyses and psychophysical investigations, there has been a growing effort in recent years to identify the neural mechanisms that are involved in the processing of 3D shape. The first sources of evidence relating to this topic were obtained from lesion studies in monkeys [35,36]. The results revealed that animals with bilateral ablations of the inferior temporal cortex are severely impaired in their ability to discriminate complex 2D patterns or shapes. Animals with lesions in the parietal cortex, by contrast, exhibit normal shape discrimination, but are impaired in their ability to localize objects in space. These findings led to a widely accepted conclusion that the primate visual system contains two functionally distinct visual pathways: a ventral 'what' pathway directed towards the temporal lobe that is involved

in object recognition, and a dorsal 'where' pathway directed towards the parietal lobe that is involved in spatial localization and the visual control of action.

The best available method for studying the neural processing of 3D shape in humans involves functional magnetic resonance imaging (fMRI), which measures local variations in blood flow in different regions of the cortex, thus providing an indirect measure of neural activation. This technique is most often used to compare patterns of brain activation in different experimental conditions. For example, to identify the neural mechanisms involved in the processing of 3D structure from motion, several investigators have compared the activation patterns produced when observers view 3D objects defined by motion relative to those that are produced by moving 2D patterns [37–40]. One limitation of this approach, however, is that it can be difficult to distinguish which specific stimulus attributes are responsible for any observed differences in neural activation. Increased activation in the 3D motion condition could be due to the processing of 3D shape, or it could be due to the processing of 3D motion trajectories. The best way of overcoming this difficulty is to compare the activation patterns for different response tasks applied to identical sets of stimuli [41,42]. Areas involved in the processing of 3D shape would be expected to become more active when making judgments about 3D shape than would otherwise be the case for other possible response tasks, such as judgments of surface texture or motion trajectories.

Recent research using both of these procedures for the perception of 3D shape from motion, shading and texture has produced a growing body of evidence that judgments of 3D shape involve both the dorsal and ventral pathways [37–40,43,44], which is somewhat surprising given the

¹ Supplementary data is available in the online version of this paper.

Box 2. Questions for future research

- Why are observers insensitive to potential information from motion and/or binocular disparity, by which it would be theoretically possible to perceive very accurately metrical relations among distance intervals in different direction? One speculative answer is that these relations might not be important for tasks that are crucial for survival.
- How do observers achieve stable perceptions of 3D structure from visual information that is inherently ambiguous? One possibility is that they exploit regularities of the natural environment to select an interpretation that is statistically most likely, although there is little hard evidence to support that hypothesis.
- How is the phenomenal experience of perceptual constancy achieved, even though observers are unable to make accurate judgments of 3D metric structure? That constancy is possible could indicate that the perceptual representation of 3D shape involves a relatively abstract data structure based on qualitative surface properties that can be reliably determined from visual information.
- How do observers identify different types of image features, such as occlusion contours, cast shadows, specular highlights or variations in surface reflectance? This is one of the oldest problems in computational vision, but researchers have made surprisingly little headway.
- How do observers obtain useful information about 3D shape from optical deformations of image features, such as smooth occlusion contours, diffuse shading, specular highlights or cast shadows, which do not remain projectively attached to fixed locations on an object's surface, as is required by current computational models of 3D structure from motion?
- What are the neural mechanisms by which 3D shapes are perceptually analyzed? Current research has identified several anatomical locations of 3D shape processing, but the precise computational functions performed in those regions have yet to be determined.

functional roles that have traditionally been attributed to these pathways. As would be expected from the results of earlier lesion studies, judgments of 3D shape produce significant activations in ventral regions of the cortex, although it is interesting to note that these do not overlap perfectly with regions involved in the analysis of 2D shape [45,46]. The analysis of 3D shape also occurs at numerous locations within the dorsal pathway, including the medial temporal cortex, and at several sites along the intraparietal sulcus.

Some of these findings are also consistent with the results obtained using electrophysiological recordings of single neurons within the dorsal and ventral pathways of monkeys [47–50]. For example, Janssen, Vogels and Orban [51–54] have shown that neurons within the inferior temporal cortex that are selective to 3D surface curvature from binocular disparity are concentrated in a small area of the superior temporal sulcus, but that neurons selective to 2D shape are more broadly distributed.

It has generally been assumed within the field of neurophysiology that the monkey visual system provides an adequate model of human brain function, but, until quite recently, there has been no way to test the validity of that generalization. That situation has changed, however, owing to recent methodological innovations [55,56], which now make it possible to perform fMRI on alert behaving monkeys using exactly the same experimental protocols as are used with humans. In one of the first studies to exploit this approach, Vanduffel *et al.* [56] compared the patterns of activation produced by 2D and 3D motion displays in humans and monkeys. Although the activations of ventral cortex were quite similar in both species, the results obtained in parietal cortex were remarkably different. Whereas the perception of 3D structure from motion produces numerous activations in humans along the intraparietal sulcus, those activations are completely absent in monkeys. These findings suggest that there may be substantial differences between the human and monkey visual systems in how visual information is analyzed for the determination of 3D shape. Because this is such a new area of research, there is much too little data at present to draw any firm conclusions. However,

this is likely to be a particularly active topic of investigation over the next several years (see also Box 2).

Conclusion

Psychophysical investigations have revealed that observers' judgments about 3D shape are often systematically distorted, but that these distortions are constrained to a limited set of transformations in a manner that is consistent with current computational analyses. These findings suggest that the perceptual representation of 3D shape is likely to be primarily based on qualitative aspects of 3D structure that can be determined reliably from visual information. One possible form of data structure for representing these qualitative properties involves arrangements of salient image features, such as occlusion contours or edges of high curvature, whose topological structures remain relatively stable over viewing directions. Other recent empirical studies have shown that the neural processing of 3D shape is broadly distributed throughout the ventral and dorsal visual pathways, suggesting that processes in both of these pathways are of fundamental importance to human perception and cognition.

Acknowledgements

The preparation of this manuscript was supported by grants from NIH (R01-EY12432) and NSF (BCS-0079277).

References

- 1 Gibson, J.J. (1950) *The Perception of the Visual World*, Houghton Mifflin
- 2 Malik, J. and Rosenholtz, R. (1997) Computing local surface orientation and shape from texture for curved surfaces. *Int. J. Comput. Vis.* 23, 149–168
- 3 Tse, P.U. (2002) A contour propagation approach to surface filling-in and volume formation. *Psychol. Rev.* 109, 91–115
- 4 Belhumeur, P.N. *et al.* (1999) The bas-relief ambiguity. *Int. J. Comput. Vis.* 35, 33–44
- 5 Koenderink, J.J. and van Doorn, A.J. (1991) Affine structure from motion. *J. Opt. Soc. Am. A* 8, 377–385
- 6 Ullman, S. (1979) *The Interpretation of Visual Motion*, MIT Press
- 7 Longuet-Higgins, H.C. (1981) A computer algorithm for reconstructing a scene from two projections. *Nature* 293, 133–135
- 8 Richards, W.A. (1985) Structure from stereo and motion. *J. Opt. Soc. Am. A* 2, 343–349
- 9 Todd, J.T. and Norman, J.F. (2003) The visual perception of 3-D shape

- from multiple cues: are observers capable of perceiving metric structure? *Percept. Psychophys.* 65, 31–47
- 10 Hecht, H. *et al.* (1999) Compression of visual space in natural scenes and in their photographic counterparts. *Percept. Psychophys.* 61, 1269–1286
 - 11 Koenderink, J.J. *et al.* (2000) Direct measurement of the curvature of visual space. *Perception* 29, 69–80
 - 12 Koenderink, J.J. *et al.* (2002) Papus in optical space. *Percept. Psychophys.* 64, 380–391
 - 13 Loomis, J.M. and Philbeck, J.W. (1999) Is the anisotropy of perceived 3-D shape invariant across scale? *Percept. Psychophys.* 61, 397–402
 - 14 Norman, J.F. *et al.* (1996) The visual perception of 3D length. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 173–186
 - 15 Koenderink, J.J. *et al.* (1996) Surface range and attitude probing in stereoscopically presented dynamic scenes. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 869–878
 - 16 Koenderink, J.J. *et al.* (1996) Pictorial surface attitude and local depth comparisons. *Percept. Psychophys.* 58, 163–173
 - 17 Koenderink, J.J. *et al.* (1997) The visual contour in depth. *Percept. Psychophys.* 59, 828–838
 - 18 Koenderink, J.J. *et al.* (2001) Ambiguity and the ‘mental eye’ in pictorial relief. *Perception* 30, 431–448
 - 19 Todd, J.T. *et al.* (1996) Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 695–706
 - 20 Todd, J.T. *et al.* (2004) Perception of doubly curved surfaces from anisotropic textures. *Psychol. Sci.* 15, 40–46
 - 21 Cornelis, E.V.K. *et al.* Mirror reflecting a picture of an object: what happens to the shape percept? *Percept. Psychophys.* (in press)
 - 22 Todd, J.T. *et al.* (1997) Effects of texture, illumination and surface reflectance on stereoscopic shape perception. *Perception* 26, 807–822
 - 23 Gibson, J.J. (1950) The perception of visual surfaces. *Am. J. Psychol.* 63, 367–384
 - 24 Gibson, J.J. (1979) *The Ecological Approach to Visual Perception*, Haughton Mifflin
 - 25 Koenderink, J.J. (1984) What does the occluding contour tell us about solid shape? *Perception* 13, 321–330
 - 26 Hoffman, D.D. and Richards, W.A. (1984) Parts of recognition. *Cognition* 18, 65–96
 - 27 Siddiqi, K. *et al.* (1996) Parts of visual form: psychological aspects. *Perception* 25, 399–424
 - 28 Singh, M. and Hoffman, D.D. (1999) Completing visual contours: the relationship between relatability and minimizing inflections. *Percept. Psychophys.* 61, 943–951
 - 29 Biederman, I. (1987) Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* 94, 115–147
 - 30 Malik, J. (1987) Interpreting line drawings of curved objects. *Int. J. Comput. Vis.* 1, 73–103
 - 31 Huffman, D.A. (1977) Realizable configurations of lines in pictures of polyhedra. *Machine Intelligence* 8, 493–509
 - 32 Mackworth, A.K. (1973) Interpreting pictures of polyhedral scenes. *Artif. Intell.* 4, 121–137
 - 33 Hummel, J.E. and Biederman, I. (1992) Dynamic binding in a neural network for shape recognition. *Psychol. Rev.* 99, 480–517
 - 34 Norman, J.F. *et al.* Perception of 3D shape from specular highlights and deformations of shading. *Psychol. Sci.* (in press)
 - 35 Dean, P. (1976) Effects of inferotemporal lesions on the behavior of monkeys. *Psychol. Bull.* 83, 41–71
 - 36 Ungerleider, L.G. and Mishkin, M. (1982) Two cortical visual systems. In *Analysis of Visual Behavior* (Ingle, D.J. *et al.*, eds), pp. 549–586, MIT press
 - 37 Kriegeskorte, N. *et al.* (2003) Human cortical object recognition from a visual motion flowfield. *J. Neurosci.* 23, 1451–1463
 - 38 Murray, S.O. *et al.* (2003) Processing shape, motion and three-dimensional shape-from-motion in the human cortex. *Cereb. Cortex* 13, 508–516
 - 39 Orban, G.A. *et al.* (1999) Human cortical regions involved in extracting depth from motion. *Neuron* 24, 929–940
 - 40 Paradis, A.L. *et al.* (2000) Visual perception of motion and 3-D structure from motion: an fMRI study. *Cereb. Cortex* 10, 772–783
 - 41 Corbetta, M. *et al.* (1991) Selective and divided attention during visual discriminations of shape, color, and speed: functional anatomy by positron emission tomography. *J. Neurosci.* 11, 2383–2402
 - 42 Peuskens, H. *et al.* Attention to 3D shape, 3D motion and texture in 3D structure from motion displays. *J. Cogn. Neurosci.* (in press)
 - 43 Shikata, E. *et al.* (2001) Surface orientation discrimination activates caudal and anterior intraparietal sulcus in humans: an event-related fMRI study. *J. Neurophysiol.* 85, 1309–1314
 - 44 Taira, M. *et al.* (2001) Cortical areas related to attention to 3D surface structures based on shading: an fMRI study. *Neuroimage* 14, 959–966
 - 45 Kourtzi, Z. and Kanwisher, N. (2000) Cortical regions involved in perceiving object shape. *J. Neurosci.* 20, 3310–3318
 - 46 Kourtzi, Z. and Kanwisher, N. (2001) Representation of perceived object shape by the human lateral occipital complex. *Science* 293, 1506–1509
 - 47 Taira, M. *et al.* (2000) Parietal neurons represent surface orientation from the gradient of binocular disparity. *J. Neurophysiol.* 83, 3140–3146
 - 48 Tsutsui, K. *et al.* (2001) Integration of perspective and disparity cues in surface-orientation-selective neurons of area CIP. *J. Neurophysiol.* 86, 2856–2867
 - 49 Tsutsui, K. *et al.* (2002) Neural correlates for perception of 3D surface orientation from texture gradient. *Science* 298, 409–412
 - 50 Xiao, D.K. *et al.* (1997) Selectivity of macaque MT/V5 neurons for surface orientation in depth specified by motion. *Eur. J. Neurosci.* 9, 956–964
 - 51 Janssen, P. *et al.* (1999) Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proc. Natl. Acad. Sci. U. S. A.* 96, 8217–8222
 - 52 Janssen, P. *et al.* (2000) Three-dimensional shape coding in inferior temporal cortex. *Neuron* 27, 385–397
 - 53 Janssen, P. *et al.* (2000) Selectivity for 3D shape that reveals distinct areas within macaque inferior temporal cortex. *Science* 288, 2054–2056
 - 54 Janssen, P. *et al.* (2001) Macaque inferior temporal neurons are selective for three-dimensional boundaries and surfaces. *J. Neurosci.* 21, 9419–9429
 - 55 Orban, G.A. *et al.* (2003) Similarities and differences in motion processing between the human and macaque brain: evidence from fMRI. *Neuropsychologia* 41, 1757–1768
 - 56 Vanduffel, W. *et al.* (2002) Extracting 3D from motion: differences in human and monkey intraparietal cortex. *Science* 298, 413–415
 - 57 Cipollo, R. and Giblin, P. (1999) *Visual Motion of Curves and Surfaces*, Cambridge University Press
 - 58 Tarr, M.J. and Kriegman, D.J. (2001) What defines a view? *Vision Res.* 41, 1981–2004
 - 59 Phillips, F. *et al.* (2003) Perceptual representation of visible surfaces. *Percept. Psychophys.* 65, 747–762

Voice your Opinion in TICS

The pages of *Trends in Cognitive Sciences* provide a unique forum for debate for all cognitive scientists. Our Opinion section features articles that present a personal viewpoint on a research topic, and the Letters page welcomes responses to any of the articles in previous issues. If you would like to respond to the issues raised in this month's *Trends in Cognitive Sciences* or, alternatively, if you think there are topics that should be featured in the Opinion section, please write to:

The Editor, Trends in Cognitive Sciences, 84 Theobald's Road, London, UK WC1X 8RR,
or e-mail: tics@current-trends.com