

# Fechner, information, and shape perception

Joseph S. Lappin · J. Farley Norman · Flip Phillips

Published online: 31 August 2011  
© Psychonomic Society, Inc. 2011

**Abstract** How do retinal images lead to perceived environmental objects? Vision involves a series of spatial and material transformations—from environmental objects to retinal images, to neurophysiological patterns, and finally to perceptual experience and action. A rationale for understanding functional relations among these physically different systems occurred to Gustav Fechner: *Differences* in sensation correspond to differences in physical stimulation. The concept of information is similar: Relationships in one system may correspond to, and thus represent, those in another. Criteria for identifying and evaluating information include (a) *resolution*, or the precision of correspondence; (b) *uncertainty* about which input (output) produced a given output (input); and (c) *invariance*, or the preservation of correspondence under transformations of input and output. We apply this framework to psychophysical evidence to identify visual information for perceiving surfaces. The elementary spatial structure shared by objects and images is the second-order differential structure of *local surface shape*. Experiments have shown that human vision is

directly sensitive to this higher-order spatial information from interimage disparities (stereopsis and motion parallax), boundary contours, texture, shading, and combined variables. Psychophysical evidence contradicts other common ideas about retinal information for spatial vision and object perception.

**Keywords** Fechner · Information · Perception · Shape · Surfaces · Space · Depth · 3-D perception · Stereo · Motion · Contours · Texture · Shading · Cue combination · Spatial vision

## Part 1. Information and the representation of structure

### Mind from matter

*The fundamental problem of vision is to understand how retinal images lead to perceived environmental objects*

Spatial vision is so marvelously effective that usually it is taken for granted. We perceive environmental objects but not the visual instrument by which they are revealed. Spatial vision pervades our conscious experience, but our understanding of the underlying visual mechanisms has substantial gaps.

Gaps in understanding occur especially at the interfaces between visual subsystems, where information is transferred from one physical and spatial format to another. How does vision transcend vast changes in format—involving 3-D objects and spaces, 2-D optical images, neural spike trains and synaptic interactions in multiple neural areas, conscious perceptions, and behavioral actions? How do immaterial knowledge and experience arise from material objects and events?

---

J. S. Lappin  
Vanderbilt University,  
Nashville, TN, USA

J. F. Norman  
Western Kentucky University,  
Bowling Green, KY, USA

F. Phillips  
Skidmore College,  
Saratoga Springs, NY, USA

J. S. Lappin (✉)  
5083 Hanging Moss Lane,  
Sarasota, FL 34238, USA  
e-mail: joe.lappin@vanderbilt.edu

The present article offers a theoretical framework for addressing these issues—based on the concept of information, involving corresponding relational structures in physically separate domains. We describe criteria for identifying and evaluating structural correspondence and then apply these criteria to research on shape perception.

### *Fechner's insight*

In the 19th century, Gustav Fechner struggled for years to find a scientific rationale for linking the material and mental worlds. Material objects have mass and spatial extent, but mental experience has neither. Some have concluded that mental events belong to a spiritual world. Fechner, however, believed that mind and body are inseparable and that mental events are emergent properties of complex material systems, such as brains (see Heidelberger, 2004, pp.118–119). His insight was that sensory discrimination experiments offer meaningful quantitative bridges from physics to psychology.

According to Fechner, the idea that came to him on “October 22, 1850 at dawn in bed” was to make “the proportionate increase in living energy ... be the measure of the increase of pertinent mental intensity” (Vol.2: Fechner, 1860, 2004, p.56; translated from Fechner, 1860, p.554). The idea was expressed mathematically in his “fundamental formula”:

$$dg = K db/b, \quad (1)$$

where  $dg$  is a change in sensation magnitude,  $db$  is a change in physical stimulus magnitude,  $b$  is a positive ratio-scaled physical magnitude, and  $K$  is an arbitrary constant. As Fechner said, “the fundamental formula ... simply expresses the relation holding between small relative stimulus increments and sensation increments. In short, it is nothing more than Weber's law...” (Vol.2: Fechner, 1860; 1965, p.70; translated from Fechner, 1860, p.10).

Discussions of Fechner's psychophysics usually focus on his logarithmic scale of sensory magnitude, but his fundamental formula is more important. Integrating the fundamental formula yields the logarithmic scale.

Fechner's insight entails two main ideas: First, *differences* in sensation can meaningfully correspond to *differences* in physical energy. Physical quantities do not meaningfully correspond to experiential quantities, but *relations* among physical quantities may correspond to relations among sensations.

A second aspect of Fechner's insight is “Weber's law”—that a detectable change in physical intensity is (approximately) proportional to that intensity. In Fechner's words, “Weber's Law ... should, because of its wide generality and the extent of the boundaries within which it is strictly or approximately valid, be looked upon as fundamental to

psychic measurement theory” (Vol.1: 1860, 1966, pp.54–55). The validity of the hypothesized invariance of the Weber fraction is not important here. The important idea is that physical and mental variations share a common structure. The Weber fraction is but one possibility.

### *Information, communication, and structural representation*

About a century after Fechner's insight, related ideas emerged in theories of information, communication, and control—generalizing the idea that relationships in one domain can correspond to and represent those in another.

Information, as the term itself suggests, is “in the formation”—based on organization rather than energy (Wiener, 1954, 1961).<sup>1</sup> Information is transmitted by corresponding variations in separate systems. Information, communication, and control involve changes, and changes involve energy, but the quantity of information is unrelated to the quantity of energy. Physical events involve transfers of energy, but communication and perception involve transmission of information. Energy is “conserved” (constant), but

<sup>1</sup> *Information* is often understood as a statistical concept, based on the probabilities of alternative states of a system, as in Shannon's (1948/1949) theory. Applications of this statistical theory to experimental psychology have been reviewed by Garner (1962) and Laming (2010). Shannon's goal was to quantify transmitted information—when the sets of messages and physical signals are already specified.

The present goal is to define information when the relevant variables are not known prior to investigation. Specifically, we seek to clarify the representation of variations in one system by those in another. The problem of representation logically precedes a statistical description.

Shannon (1948/1949) sought a general quantitative method applicable to any variable, regardless of its physical dimensions or scale, applicable to symbolic as well as physical variations. Accordingly, variations were treated simply as categories, where the relations among alternative values are simply *same* or *different*, members of the same set or different sets.

An expanded set of relationships is required to understand representation more generally, especially those involving variations found in nature. Optical and spatial variations relevant to visual shape perception, for example, involve ordinal and topological structures in multiple dimensions of space and time and involve differentials of at least second order.

Shannon's generalized representation of categorical structures involved no assumptions about physical form or relations among categories. Categorical structures are readily represented by sets of symbols, but symbols are less useful for many other relational structures. A common misunderstanding is that “information” is fundamentally symbolic, without physical or geometric or meaningful structure. This unfortunate misconception fails to recognize the great variety of forms of information in physical and biological systems. It is also insufficient for characterizing representation in general.

The informal concept of ‘information’ used in both science and everyday language implicitly encompasses a wide range of physical forms and relationships. The present aim is to develop a generalized concept of information relevant to visual perception and other applications.

information is not conserved and not governed by natural laws of space, time, mass, and energy.

To understand vision, one must identify the information it uses. Retinal images and physiological responses can lead to perceived environmental objects if and only if these physically different variables have a corresponding structure. As Marr (1982) pointed out, “a process may be thought of as a mapping from one representation to another” (p.31). Thus, representations of input and output information shape inferences about visual or computational processes for mapping one to the other. Identifying *what* information is processed is a prerequisite for determining *how* it is processed.

Marr (1982, p.31) represented optical input as a scalar field of intensity values, spatially structured by retinal anatomy. Marr and many others have regarded this representation as obvious, but it is a *hypothesis*, one of several empirically testable alternatives. Abundant experimental evidence that will be reviewed in Part 2 of this article shows that retinal information has topological structure—involving relations among image points.

Many of the present ideas resemble those of James Gibson. Gibson (1950) concluded that visual information is based on invariant higher-order spatiotemporal structure, and we seek to identify that structure. Surface structure had a basic role in Gibson’s conception of sensory information, and it is also basic here. Gibson (1966, 1979) also emphasized that (a)the structure of sensory stimulation reflects the structure of the environment, and (b)relevant environmental structure covaries with the perceiver’s perspective, actions, and attention. Those ideas are also reflected here. Not surprisingly, rationale for these ideas may be found outside Gibson’s writings.

The present approach generalizes concepts of information beyond symbolic and statistical structures (e.g., Laming, 2010; Shannon, 1948/1949) to encompass geometric organization. This generalization is not new. Applications of cybernetics to biology (Ashby, 1963; Meyer-Eppler, 1969; Wiener, 1961) led naturally to descriptions of sensory information as spatiotemporal patterns structured by interactions with the environment.

### *Psychophysics and representation*

The principal goal of psychophysics is sometimes regarded as the numerical representation of sensations; that was a primary aim of Fechner, and has been of many others since. Luce (2004) and Steingrímsson (2009) are contemporary examples. The present application, however, entails nonnumerical representation.

The first problem of psychophysics is *what* to measure. As S.S. Stevens put it, “In a sense, there is only one problem of psychophysics, namely the definition of the

stimulus. In this sense, there is only one problem in all of psychology—and it is the same problem” (1951, p.31). If “stimulus” is replaced by “information,” then Stevens states the present problem: What is the information for perceiving environmental objects?

In general, representation involves a correspondence between two relational structures.<sup>2</sup> The theoretical rationale for representation has been studied in the theory of measurement, concerned with the problem of representing an empirical structure of qualitative relations by a structure of numerical relations (e.g., Krantz, Luce, Suppes, & Tversky, 1971; Roberts, 1979; Suppes & Zinnes, 1963). Two subtasks in measurement are known as the “representation” and “uniqueness” problems.

Both problems have analogues in studying biological systems and in identifying optical and visual representations of environmental objects. Methods for identifying such representations are developed in the section below on “Evaluating structural correspondence”; the methods are then illustrated by application to the specific problem of shape perception in Part 2 of this article.<sup>3</sup>

### **Basic concepts and terminology**

Many different terms might be used for concepts of structure and information. We will rely on the following terminology:

*Forms* are variable local elements that carry information. For shape perception, the local forms are spatial variables that occur in smoothly varying arrays on surfaces and their images.

“Form” is used here as a generic term for both spatial and nonspatial elements. Symbolic forms—for instance, letters, words—label discrete categories, where spatial shape is arbitrary except to distinguish among alternatives. Temporal forms occur in music, speech, and other acoustic events. In spatial vision, the elementary forms of moving objects deform smoothly over time in moving images and change discretely between different viewpoints and different objects.

*Structure* refers here to relations among forms—synonymous with the concept of “relational structure” in measurement theory.

<sup>2</sup> A *relational structure* is a set of elements, relations among the elements, and (often) operations that yield other elements in the set.

<sup>3</sup> *Representation* refers here to a correspondence between physically separate relational structures—where variations in one system correspond to, and thus represent, those in another. The same term is applied to many different types of representation, including spatiotemporal, numeric, and symbolic structures, natural or biological structures, as well as artificial structures of human design. The term implies neither mental nor symbolic representation.

Corresponding structures do not imply corresponding forms. The elementary forms typically differ between domains—as in linguistic and numerical representations—and in visual correspondences between objects, images, physiological responses, and perceived objects. Correspondence involves the relational structure, not the forms themselves.

Here, *information* is said to exist if and only if two or more physically distinct structures correspond to one another. “Information” is distinguished from the concrete forms or signals that carry it. Thus, “information” is neither objectively defined nor concrete, because (a) it involves relationships among forms rather than forms per se, (b) relational structure may be represented in multiple ways, and (c) it requires corresponding representations by both receiver and sender. Despite its nonobjectivity, this approach resolves basic ambiguities in identifying information, because arbitrary physical variables seldom correspond across systems. In Shannon’s (1948/1949) statistical theory of communication, the structures were defined a priori, but in sensory systems, the information-carrying forms and structures are generally not known prior to investigation.

### Evaluating structural correspondence

#### *Spatial correspondences of objects, images, and perceptions*

Structural organization can be represented in many ways. No structure can be singled out, especially beforehand, as objectively correct or optimum. How, then, can relevant, information-carrying structure be identified?

Three quantifiable criteria for evaluating structural correspondence are *resolution*, *uncertainty*, and *invariance*. The aim is to identify corresponding relational structures that maximize resolution, minimize uncertainty, and maintain correspondence under transformations of observational conditions. Correspondence of particular relational structures may be quantified by Shannon’s (1948/1949) information theory.

Three physical domains are relevant: environmental objects, optical images, and physiological patterns—say  $O$ ,  $I$ , and  $\Psi$ . The relevant structures in these domains are inferred (by a scientist) from correspondences across domains. Structures in  $O$  and  $I$  are scientifically observable but permit multiple representations. Structures in  $\Psi$  are not directly observable by scientists, and can only be inferred.<sup>4</sup>

<sup>4</sup> The structure of  $\Psi$  correlates with patterns of neural spike trains and synaptic interactions. Presently available scientific observations do not reveal the information-carrying organization of these physiological patterns, however. Instead, physiological structure must be inferred from correspondences among objects, images, and behavioral discriminations.

The maps between structures in these domains may be labeled  $f$ ,  $g$ , and  $h$ , along with their inverses:

$$\begin{aligned} f : O &\rightarrow I & f^{-1} : I &\rightarrow O', \\ g : I &\rightarrow \Psi & g^{-1} : \Psi &\rightarrow I', \text{ and} \\ h : \Psi &\rightarrow O'. \end{aligned}$$

Three different sets of the environmental objects— $O$ ,  $O'$ , and  $O''$ —are distinguished because these are defined differently. The set of environmental objects,  $O$ , may be described by a scientist without reference to retinal or photographic images, but another set of objects,  $O'$ , indirectly defined by their images, is typically smaller, including multiple environmental objects with the same or indistinguishable images. Similarly, a third set of perceived objects,  $O''$ , is operationally defined by behavioral discriminations. A distinction between the image structures  $I$  and  $I'$  is also needed, since  $I'$  is defined by behavioral discriminations. The inverse map  $h^{-1}$  is not meaningful in this context and is safely ignored. Combining functions, the map from physical to perceived objects is

$$h g f : O \rightarrow O''.$$

These are generic maps that might not satisfy correspondence criteria and might differ in dimensionality. Structures and forms that carry information are subsets of many possible representations. Identifying information requires investigation. Figure 1 illustrates these maps.<sup>5</sup>

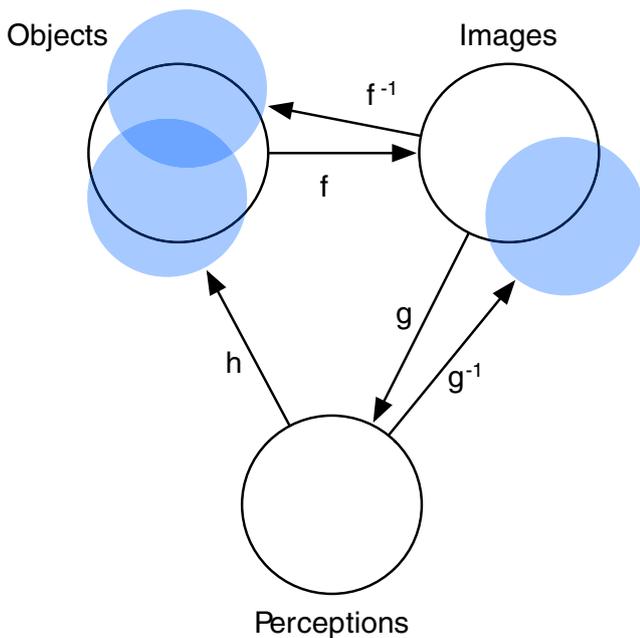
#### *Resolution*

*Resolution* refers to the precision of representation. *Bandwidth* and *capacity* are related concepts involving the complexity of a representation. In psychophysical experiments, resolution is often evaluated by discriminations among neighboring signals, but dynamic range is also relevant.

Shannon’s (1948/1949) measure of information transmission offers a general method of quantifying resolution. This measure assumes nothing about the structure of relations among the alternative forms. In Shannon’s theory,

<sup>5</sup> These functions map variations—structures of possible elementary forms—not individual forms, symbols, or objects. Transfers of information involve sets of potential events, not particular causal events or processes operating on individual stimuli.

The sets of objects, images, and perceived objects in Figure 1 involve an investigator’s representations of forms and structures of variation among these forms. Thus, the representations may be adjusted to maximize correspondence across the domains of environment, optical images, and behavioral discriminations. The quantitative evaluations of uncertainty and information transmission—for instance,  $H(X)$  and  $T(X:Y)$  as defined in the next section—are contingent on the representations of these sets of objects, images, and perceived objects.



**Fig. 1** Visual information involves structural correspondences between multiple physical domains: environmental objects, retinal images, and perceptions. Objects defined by the inverse map from images to objects,  $f^{-1}: I \rightarrow O'$ , differ from the input environmental objects,  $O \neq O'$ , and a third set of objects,  $O''$ , is defined by behavioral discriminations,  $h: \Psi \rightarrow O''$ . Similarly, behaviorally discriminated images,  $g^{-1}: \Psi \rightarrow I'$ , also differ from the input retinal images,  $I \neq I'$ . Psychophysical analyses of visual information transmission are based on empirical correspondences such as  $T(O:O'')$

the structure is simply a set of categories, with relations *same* versus *different*. Thus, the variability of a set  $X$  of categories,  $x_i \in X$ , is given as

$$H(X) = - \sum \{p(x_i) \log_2 [p(x_i)]\}, \tag{2}$$

where  $p(x_i)$  is the probability of occurrence of category  $x_i$ , with  $\sum [p(x_i)] = 1.0$ . The probability distribution is not important in most of the present geometric analysis, and in this article, the categories usually are treated as equally likely,  $p(x_i) = p(x_j)$ .

The information transferred by corresponding variations of two different sets  $X$  and  $Y$  is given by

$$T(X:Y) = H(Y) - H(Y|X), \tag{3}$$

where  $H(Y|X)$  is the conditional uncertainty of  $Y$  given  $X$ . Thus, if  $H(Y|X) = 0.0$ , then  $T(X:Y) = H(Y)$ ; if  $H(Y|X) = H(Y)$ , then  $X$  and  $Y$  are statistically independent, and  $T(X:Y) = 0.0$ . This correspondence measure is symmetric,  $T(X:Y) = T(Y:X)$ . The maximum value of  $T(X:Y)$  is the smaller variability of the two sets:  $T(X:Y) \leq \min\{H(X), H(Y)\}$ .

Psychophysical experiments cannot directly access forms or structures in  $\Psi$ , so we rely instead on observable

correspondences such as  $T(O:O'')$ . “Low-level” visual resolution of spatial structure may be evaluated as  $T(I:I')$ , and perceptual discriminations of environmental objects may be evaluated as  $T(O:O'')$ . One aim of psychophysical experiments is to discover what vision sees best—that is, what spatial structures maximize  $T(I:I')$  and  $T(O:O'')$ .

*Uncertainty*

*Uncertainty* refers in this article to the set of forms corresponding to a given form in another structure. When a map from one structure to another is many-to-one, the inverse map is uncertain. Uncertainty implies ambiguity. In information theory terms, the present use of “uncertainty” refers to conditional uncertainty—for instance,  $H(O'|I)$ , the set of objects that potentially correspond to a given image. Image ambiguity is basic in analyzing spatial vision.

Geometric uncertainties are especially relevant. Projected images obviously provide information about 3-D spatial structure, but images have inherent ambiguities. Perceptual uncertainties,  $H(O''|\Psi)$ , often arise from image uncertainties,  $H(O'|I)$ . Environmental and perceived objects necessarily correspond less than objects and images,  $T(O:O'') < T(O:O')$ , but the differences can be small.

*Invariance*

*Invariance* refers to the permissible transformations of a structure that do not alter its correspondence with another structure. Indeed, “structure” is defined by the groups of transformations that leave it invariant. Thus, invariance is a fundamental aspect of the concept of information, and it is fundamental in identifying information in natural systems. The set of potential information-carrying structures is sharply reduced by the requirement that visual resolution is maintained under a particular group of transformations.

Perceptual constancies of an object’s shape under observational transformations—such as changes in viewpoint, illumination, and so forth—are not just puzzles awaiting explanation; they define perceived “shape.” The invariance criterion is essential in experimental research on shape perception. The experimental aim is to test whether discriminations between objects are robust under changes in viewpoint, for example.<sup>6</sup>

In theoretical physics and mathematics, invariance is often called *symmetry*—equivalence of structure after

<sup>6</sup> Consider the map  $m=hgf$  from environmental objects to perceived objects,  $m: O \rightarrow O''$ . Suppose that  $\tau: O \rightarrow O$  is a group of transformations that map objects in  $O$  onto other objects in  $O$ . If  $T[\tau(O):O''] \approx T(O:O'')$ , then we say that the correspondence between  $O$  and  $O''$  is *invariant* under the transformation group  $\tau$ . This is an experimentally testable criterion. The method is especially powerful if the transformations are randomly varied, where  $H[\tau(O)] >> H(O)$ .

specific groups of transformations. All physical laws may be expressed as symmetries—as structural relations that are conserved under specific groups of transformations of observational parameters (Lederman & Hill, 2004). Conservation of energy, for example, is equivalent to the invariance of physical interactions under shifts in time of occurrence. Symmetry has several (related) meanings and applications in mathematics, which generally involve equivalent structures of different mathematical objects. Symmetries between different mathematical systems, such as algebra and geometry, offer methods for revealing patterns in one structure through investigations of another (Descartes, 1637/1886; Stewart, 2007).

Invariance (symmetry) is a powerful research tool, because a structure may be *deduced* from its invariance. Rather than inducing invariance from empirical observations, one can postulate invariance under a specific transformation group and then search, theoretically and experimentally, for a structure that satisfies the invariance. Symmetry is used in this way in physics and mathematics. In measurement theory, “uniqueness” of the numerical representation specifies the scale of measurement (Luce, Krantz, Suppes, & Tversky, 1990, chaps. 20 & 22; S.S. Stevens, 1951; Suppes & Zinnes, 1963). In psychology, the methods of converging operations (Garner, Hake, & Eriksen, 1956) and the receiver operating characteristic (e.g., Swets, 1996) exemplify use of the invariance criterion, but other uses have been limited. Importantly, visual information about the shapes of environmental objects is identified by its invariance under changes in viewpoint, illumination, and context.

### Measuring resolution

Psychophysical resolution may be measured in many ways. Shannon’s statistical measure of transmitted information is general: It involves merely discrimination among alternatives and is applicable to any relational structure—symbolic, spatiotemporal, numeric, and so forth. Information often involves relations of order, geometry, or so on, and resolution measures may usefully reflect such structure. In fact, information theory measures such as  $T(X:Y)$  are rarely used in contemporary psychophysics, and were not used in the experiments reviewed in this article.

Spatial vision often involves ratio-scaled dimensions such as length. The Weber fraction,  $\Delta s/s$ , is often useful in such cases—where  $\Delta s$  is a discriminable difference in, say, spatial position, and  $s$  is a spatial length. The Weber fraction is a dimensionless measure applicable to many physical variables. The comparative visual resolution of various variables (length, depth, slant, curvature, etc.) may be evaluated by their Weber fractions. Large Weber fractions (e.g., 15% or greater) indicate poor resolution.

A similar measure is the coefficient of variation,  $SD/M$ , where  $SD$  and  $M$  are the standard deviation and mean of a positive ratio-scale variable. If one uses the standard deviation as the numerator of the Weber fraction,  $\Delta s$ , then the Weber fraction and the coefficient of variation are equivalent measures applicable to different experimental procedures.

A related approach is used in signal detection theory, where resolution is scaled in units of standard deviations. The Weber fraction and the coefficient of variation can be computed from a pair of physical values for which  $d' = 1.0$ .

Another version of the Weber fraction is the familiar *contrast ratio*,  $C = (x_1 - x_2)/(x_1 + x_2)$ , where  $x_1$  and  $x_2$  are the values of two stimuli on a positive-valued ratio-scaled dimension. If the difference in stimuli is such that discrimination accuracy,  $d'(x_1, x_2) = 2.0$ , then each of the two stimuli differs from their mean,  $(x_1 + x_2) / 2$ , by one standard deviation.

## Part 2. Visible information about surface structure

### Image information

#### Image resolution of surfaces

Given the vast spatial and material differences between objects and their optical images, what spatial commonalities support the perceptual effectiveness of images? A basic answer is that the spatial structure of images corresponds to the spatial structure of environmental *surfaces*.

Koenderink and van Doorn contributed this fundamental idea (Koenderink, 1984a, 1990; Koenderink & van Doorn 1975, 1976a, 1976b, 1976c, 1980, 1992a, 1997; Koenderink, van Doorn, Kappers, & Todd, 2001; Lappin & Craft, 2000). Key aspects include:

1. Environmental surfaces and their images are both 2-D manifolds. Both are differentiable, and both can be described by spatial derivatives.
2. Spatial relations on smoothly connected regions of a surface and on generic images of those regions are related to one another by local linear coordinate transformations. These coordinate transformations vary smoothly over the surface, with the relative orientation, curvature, and shape of the surface.<sup>7</sup>

<sup>7</sup> “Smooth” surfaces are differentiable almost everywhere. Polyhedral objects in which planar faces join at a sharp edge may be treated as smooth, since they are differentiable up to an arbitrarily small and spatially isolated region at the edge. This description of smoothly curved surfaces includes polyhedra, such as those analyzed by Pizlo (2008). Pizlo’s analysis concerned mainly global shape properties, whereas the present analysis is primarily local.

Thus, the differential structures of smooth regions of a surface and its image are approximately isomorphic (“diffeomorphic”). There is a smooth map from one manifold onto the other with a smooth inverse—say,  $f: O \rightarrow I$ , and  $f^{-1}: I \rightarrow O$ , where  $f$  and  $f^{-1}$  are both smooth. See Fig. 2. Violations of this isomorphism can result from several causes, including random noise and “accidental” violations of spatial continuity.

Nevertheless, a given image corresponds to a *group* of permissible surfaces, which share a common spatial structure but differ in relative depths (as discussed below).

3. The structure of a given surface may be defined by (a) multiple image variables—motion, binocular disparity, texture, shading, boundary contours—and (b) multiple images—from different viewpoints, focal lengths, illuminations, contexts, observer movements, and so on. Importantly, multiple images of the same surface are linearly related to the same underlying structure. Thus, adding variables and images usually reduces uncertainty about the surface (Koenderink & van Doorn, 1997).

*Spatial uncertainties about objects given their images*

Images in our eyes and in pictures yield compelling impressions of spatial structure, and images provide real-time guides for our actions. Nevertheless, images have inherent ambiguities (see Koenderink, 2001). Spatial uncertainties,  $H(O|I)$ , arise from several sources:

*Optical noise, resolution, and scale space:* At the photonic level, optical information is statistical. Light reflected from a given point on a surface is spatially distributed in the image by scattering, diffraction, and refraction.

Information about the spatial structure of images is also affected by the sampling density of photoreceptors and by neural encoding of “local signs.” These factors influence most aspects of spatial vision—for instance, acuity, peripheral vision, low vision, amblyopia, “tarachopia” (Hess, 1982),

and the apparent visual field (Koenderink, van Doorn, & Todd, 2009). Image motions add information about spatial order that partially compensates for neural limitations (Lappin, Tadin, Nyquist, & Corn, 2009).

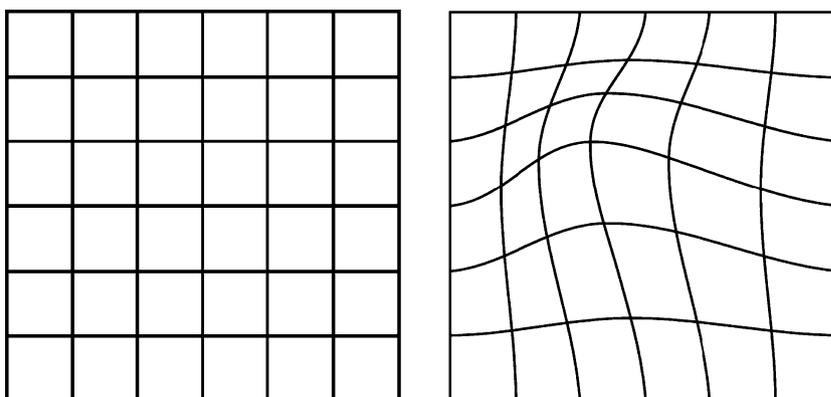
An important literature on “scale space” deals with the effect of image resolution on representations of surface structure (Florack, 1997; Koenderink, 1990). The basic idea will suffice here: *Complexity of surface topography in images should be monotonically related to image resolution* (Koenderink, 1990, chap.9). Discrete pixels can introduce high-frequency artifacts (jagged curves), but local smoothing operators can remove these artifacts.

*Spatial continuity:* Spatial continuity is potentially ambiguous in three situations.

1. Disconnected surface regions may appear connected in an image, when boundary contours of separate surface regions happen to coincide. The “Ames chair” is an illustration (Ittelson, 1952). Such “accidental views” are rare, however, and disappear with shifts in viewpoint.
2. More commonly, physically connected surface regions may be disconnected in an image, due to occlusion by a nearer surface. The relative image locations of these discontinuities change with the viewer’s position, however.
3. In experiments and occasionally in natural images, surfaces may be represented only as interpolated relations among spatially separate texture elements. Psychophysical experiments with “point-light walkers” (e.g., Blake, 1993; Fox & McDaniel, 1982; Johansson, 1973) and random-dot stereograms (Julesz, Papatomas, & Phillips, 2006) and movies (e.g., Braunstein, 1962; Lappin & Bell, 1976; Lappin & Craft, 2000; Rogers & Graham, 1979; Turner, Braunstein, & Andersen, 1995) have demonstrated that coherent image motion facilitates perception of connected structure.

*Image shading:* Changes in surface orientation cause changes in image illuminance, but changes in image

**Fig. 2** Diffeomorphic spaces. The rectilinear coordinate space on the left may be transformed into the space on the right by a smoothly varying linear coordinate transformation. As should be evident, these coordinate transformations describe local surface shape



intensity are also caused by the field of illumination, viewpoint, and surface scattering. The surface locations of image highlights are not fixed but vary with the directions of illumination (a hemispherical field) and viewpoint. Binocular highlights have different locations in each eye, and the highlights change with movements of observer and object. Only in rare cases, with Lambertian scattering, are variations in relative intensity independent of the viewpoint.

Nevertheless, the second-order differential structure of the image (associated with surface curvature in two dimensions, as described in the next section on the invariance of surface shape) tends toward invariance with illumination and viewpoint—as illumination directions are dispersed, as surface scattering increases, as viewing distance increases, and at discontinuities between surfaces (Koenderink & van Doorn, 1980). The extrema of image shading tend to occur near changes in surface shape (at parabolic lines—see Koenderink & van Doorn, 1980).

Shading in a given image region may be affected by multiple sources and surfaces—indirect reflections, shadowing, and intervening media such as glass and fog. Uncertainties about local surface structure are reduced by additional image data—wider field of view, added viewpoints, additional objects with the same illumination, surface texture, and so forth. Distinguishing the multifaceted influences of the illumination field, surface reflectance, and viewpoint is akin to solving simultaneous linear equations (Anderson & Winawer, 2008; Koenderink, Pont, van Doorn, Kappers, & Todd, 2007; Koenderink & van Doorn, 1997; Pont, 2011).

Image shading does not determine the depth scale (*bas-relief*) of the surface (Belhumeur, Kriegman, & Yuille, 1999; Koenderink et al. 2001; Todd, 2004). Stretching an object in depth does not necessarily alter its shaded image.

*Linear perspective:* The depth scale of photographs and other images depends on parameters such as focal length and the distance of the object. Without prior knowledge of global scene and object structure, image structure alone does not determine 3-D lengths, directions, angles between lines, or slants of planes. This basic image ambiguity is readily illustrated by comparing photographs of the same scene using lenses of different focal lengths: The photos are clearly different, but usually none appears either distorted or uniquely “correct.” Imaging parameters such as focal length and vantage point are ambiguous in pictures and even in movies, even though they appear subjectively unambiguous in specific cases (Cutting, 1987).

Linear perspective and depth scale are also uncertain in retinal images, despite the limited range of focal length. Many experiments have found that depth judgments are inconsistent between observers, objects, viewpoints, and sometimes even repeated observations by the same observer (e.g., Koenderink, van Doorn, & Kappers, 1996b; Lappin,

Shelton, & Rieser, 2006; Norman, Crabtree, Clayton, & Norman, 2005; Norman, Norman, Lee, Stockton, & Lappin, 2004; Norman & Todd, 1996, 1998; Norman, Todd, Perotti, & Tittle, 1996; Todd, Koenderink, van Doorn, & Kappers, 1996; Todd & Norman, 2003; Wagemans, van Doorn, & Koenderink, 2011).

*Summary:* Key aspects of image uncertainties include the following: (1) Image structure does not specify metric structure in the environment; spatial scale in depth relative to the frontal plane is indeterminate. (2) Image structure has a robust correspondence with the connectedness and topography of object surfaces. (3) Image uncertainties about environmental surface structure decrease with added image data, especially from moving observers and objects.

#### *Invariance of surface shape*

What specific aspects of surface structure reliably correspond to image structure, at least approximately invariant with changes in viewpoint and focal length? The answer: *Local surface shape*.

Simpler spatial features of surfaces, associated with relative depth (zero-order structure) and slant (first-order structure), do not satisfy the invariance criterion for information. The critical importance of local surface shape is revealed by its invariance. These ideas about invariance are illustrated in Fig. 3.

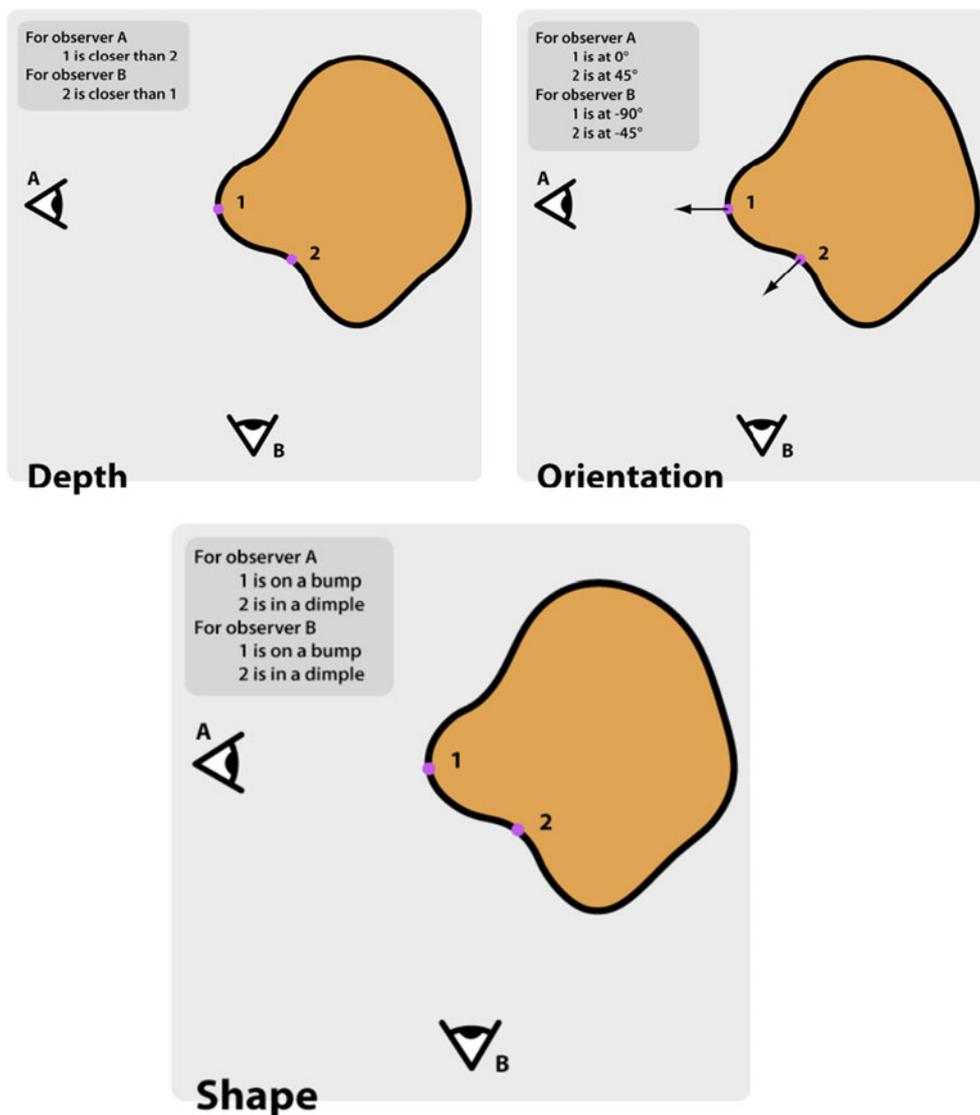
Surface shape is studied in the classical theory of differential geometry, and more recently in the psychophysical literature (see Gray, 1993; Koenderink, 1990; Koenderink & van Doorn, 1992b; Lappin & Craft, 2000; Phillips & Todd, 1996; Phillips, Todd, Koenderink, & Kappers, 2003). Here we mention only a few basic facts relevant to visual psychophysics.

At each point on a smooth (differentiable) surface, local shape is defined by the two *principal curvatures*—the maximum and minimum curvatures,  $\kappa_{\max}$  and  $\kappa_{\min}$ , which are orthogonal on the surface. Surface curvature in a given direction is given by a second-order spatial derivative, measured as  $\kappa = 1/r$ , where  $r$  is the radius of a circular arc that coincides with the surface at the given point.

Curvature in any given direction is the rate of change in direction of the surface normal relative to changes in position along the surface. This curvature is a 3-D property, not directly observable in 2-D images of the surface. Nevertheless, *the ratio of the principal curvatures is intrinsic to the surface* (i.e., is independent of a 3-D reference frame), *is defined in images of the surface, and specifies the local surface shape*.

Certain basic properties of surface shape are also given by the *Gaussian curvature*,  $K = \kappa_{\max} \kappa_{\min}$ . Surface points at

**Fig. 3** Local surface shape is invariant with changes in viewpoint. Relative depths and 3-D orientation vary with the viewpoint



which  $K = 0$ ,  $K > 0$ , or  $K < 0$  have qualitatively different shapes; they are intrinsically different and have qualitatively different images. If  $K > 0$ , where  $\kappa_{\max}$  and  $\kappa_{\min}$  have the same sign, then the surface is *elliptic*—concave or convex, valley or hill. If  $K < 0$ , where  $\kappa_{\max}$  and  $\kappa_{\min}$  have opposite signs, then the surface is *hyperbolic*—saddle-shaped. If  $K = 0$ , then the surface is either planar or parabolic; if  $\kappa_{\max} = \kappa_{\min} = 0$ , then the surface is *planar* (flat); if  $K = 0$  but  $\kappa_{\max} \neq 0$ , then the surface is *parabolic*—cylindrical, or flat in one direction but curved in others. Parabolic points occur in continuous nested curves at boundaries between elliptic and hyperbolic regions. Parabolic lines describe the surface topography. The shape at every point on a smooth surface is one of these four qualitative types (elliptic, hyperbolic, parabolic, or planar). Figure 4 illustrates the surface topography of a randomly shaped smooth solid object.

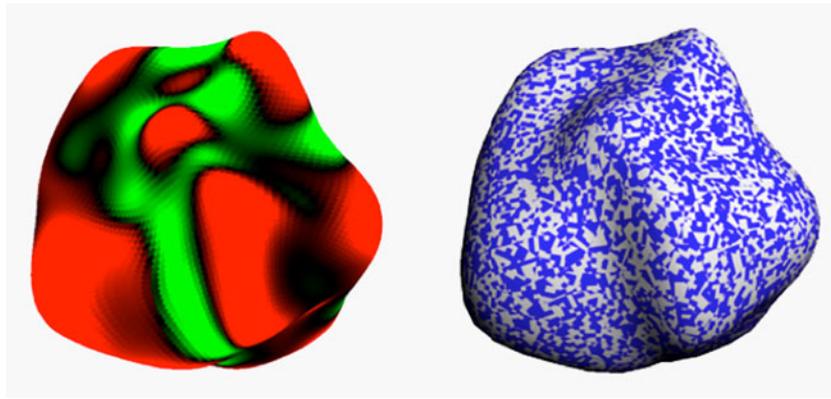
Among the possible numerical scales of shape is Koenderink and van Doorn’s (1992c; Koenderink, 1990) *shape index*,

$$S = (-2/\pi) \arctan [(\kappa_{\max} + \kappa_{\min})/(\kappa_{\max} - \kappa_{\min})]. \tag{6}$$

The shape index,  $S$ , scales shape variations by numbers in the interval  $(-1,+1)$ , as illustrated in Fig. 5. As the difference in principal curvatures,  $k_{\max} - k_{\min}$ , approaches zero, then  $S$  approaches  $+1$  or  $-1$ , where curvature is equal in all directions. This index scales qualitative surface variations that are independent of size, depth, and orientation.

The “curvedness” at each surface location is given by the combined values of the principal curvatures. Koenderink and van Doorn’s (1992c) “curvedness index” is a useful ratio scale of surface curvature:

$$C = [(k_{\max}^2 + k_{\min}^2)/2]^{1/2}. \tag{7}$$



**Fig. 4** An illustration of the surface topography described by Gaussian curvature. (Right) A randomly shaped smooth solid object, which resembles one often used in studies by Norman and Todd (e.g., Norman et al., 2006). This shape is visible by virtue of its boundary

contours, texture, and shading. (Left) Red areas are elliptic (bumps or dimples); green areas are hyperbolic (saddle-shaped); and the black areas separating the elliptic and hyperbolic are parabolic, where a principal curvature vanishes in one direction

Variations in shape and curvedness,  $S$  and  $C$ , can be pictured as polar coordinates within a Cartesian coordinate frame given by the two principal curvatures, as shown in Fig. 6. Curvedness is a 3-D property, not intrinsic to the surface, and not generally specified in images.

*Summary:* Image information about local surface shape involves the second-order differential structure of images. Local surface shape is an intrinsic surface property, independent of a 3-D coordinate frame. Image information about shape is invariant with the object's distance and orientation relative to the image. Depth relief, curvedness, and slant, however, require a 3-D reference frame. Images offer unreliable information about these 3-D properties.

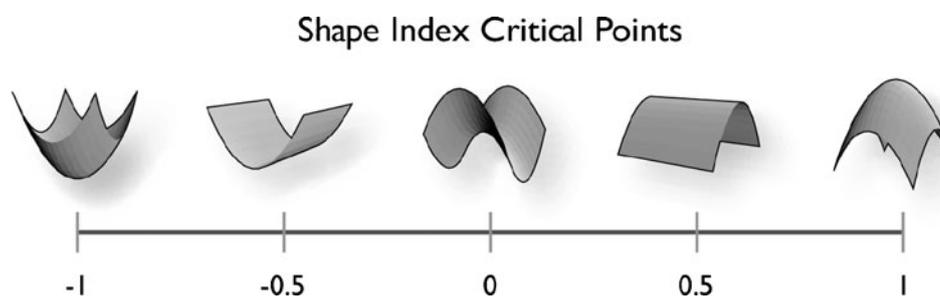
### Perception and psychophysics of surface shape

What image information supports perception of surface shape? This question may be answered by psychophys-

ical experiments, by evaluating the *resolution*, *uncertainty*, and *invariance* of human discriminations. The following psychophysical results converge with analyses of image structure to indicate that vision is sensitive to image information about local surface shape—involving the second-order differential structure of images of surfaces. Extensive psychophysical evidence is available about the roles of binocular disparity and motion parallax in spatial vision, so we focus first on this evidence.

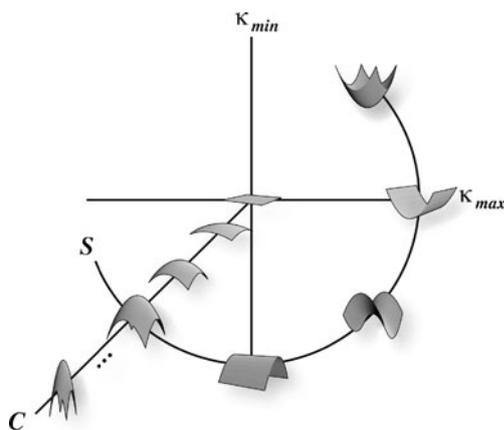
#### *Binocular disparity and motion parallax*

Stereoscopic vision derives from spatial differences (“disparities”) between two simultaneous images of an object from different viewpoints. Similar interimage disparities occur sequentially when an object and observer rotate in depth relative to one another. In both cases, the spatial disparities vary with relative depth, and both are visually effective information about surface structure. To simplify,



**Fig. 5** Local surface shapes can be described as a one-dimensional variable, scaled here by Koenderink and van Doorn's (1992c) shape index (see the text). [The illustration is from Phillips & Todd, 1996,

Fig. 3, p.932. Copyright 1996 by the American Psychological Association. Reprinted with permission.]



**Fig. 6** Koenderink and van Doorn's (1992c) shape index,  $S$ , and curvedness index,  $C$ , describe a two-dimensional space. [The illustration is from Phillips & Todd, 1996, Fig. 5, p.933. Copyright 1996 by the American Psychological Association. Reprinted with permission.]

we focus on binocular disparity, but similar comments apply to motion parallax.

What exactly is binocular disparity? The answer is less obvious than one might expect. Differences in the spatial properties of two optical images can be described in several ways, but these possibilities are seldom considered in the literature on binocular vision. The issue of defining binocular disparity exemplifies the representation problem in spatial vision research.

A common assumption is that binocular disparity is a difference between the two retinal positions of a given image feature. That is, the elementary spatial form is assumed to be a single point (e.g., an individual texture element), and its spatial positions in the two eyes are regarded as given by the anatomy of the two retinas.

An alternative class of spatial representations involves topological relations between a given point and its surrounding neighborhood. One may recognize at the outset that (a) relative spatial positions might be represented in multiple ways, involving relationships with various numbers of surrounding points, and (b) we do not know beforehand, without investigation, which of these possibilities best represents the effective optical input for binocular vision.

The experimental evidence reviewed here shows that binocular vision exhibits (a) very high resolution for detecting disparities (ignoring for the moment exactly how these are defined), (b) uncertainties about the depth scale for disparities, with corresponding uncertainties about surface slant and curvedness, and (c) invariance of perceived local surface *shape* under image transformations associated with 3-D shifts in viewpoint. This evidence implies that binocular vision is directly and especially sensitive to

second-order differential structure of disparities associated with local surface shape.<sup>8</sup>

*Resolution:* Vision is said to have “hyperacuity” for interimage disparities in spatial position (Westheimer, 1975, 1977, 1979). *Hyperacuity* is an apt term: Human discrimination thresholds for binocular disparity and relative motion are well below seeming physical limits of about 30 arcsec, based on optical diffraction, the eye’s line-spread function, and photoreceptor spacing.<sup>9</sup>

Monocular acuity for relative position is good, sometimes as low as 30 arcsec. For example, without disparity, a 2° space between two parallel vertical lines can be bisected with precision ( $SD$ ) near 30 arcsec, and the Weber fraction ( $SD$ /separation) is about 1% over varied line separations and configurations (De Valois, Lakshminarayanan, Nygaard, Schlussel, & Sladky, 1990; Lappin & Craft, 2000).

Acuity for binocular disparity is much better, however. Using the same three-line configuration, thresholds ( $SD$ ) for binocular disparity are about 10 arcsec for 2° reference-line separations, and Weber fractions are below 0.3% (e.g., Lappin & Craft, 1997, 2000). Thus, the resolution of binocular *differences* in spatial position is substantially better than spatial resolution in either of the monocular images alone! Another example: The disparity threshold is below 5 arcsec for detecting a depth difference between two adjacent planes of a 60-Hz dynamic random-dot stereogram (Cormack, Stevenson, & Schor, 1991).

Hyperacuties for relative motion are quantitatively and functionally similar to those for binocular disparity (Lappin & Craft, 2000; Rogers & Graham, 1983). These hyperacuties are robust and well documented.

*Uncertainty,* Despite hyperacuity for detecting binocular disparity, vision does not reliably represent magnitudes of

<sup>8</sup> The definition of a relationship does not require definition of the elementary components of the relationship. This is true for both mathematical and sensory representations. The Weber fraction—a dimensionless ratio—is a good example. Definitions and evaluations of derivatives are not necessarily derived from the specific values involved in a function or change. Visual representations of differential structures of optical patterns do not require elementary retinal positions of component points. Higher-order differential structure of images may be measured directly by local operators that are higher-order derivatives of Gaussians, without subtracting differences between lower-order components (see Koenderink, 1990). Indeed, measuring differences by subtracting measures of lower-order values is vulnerable to noise: The variance of a difference between two independent variables equals the sum of their two variances.

<sup>9</sup> The mechanisms underlying visual hyperacuties for binocular disparity and relative motion do not supersede physics, of course, but explanations in the literature are often unclear. Evidently, the mechanism involves spatial phase interference between coherent physiological signals from the two eyes (Lappin & Craft, 1997, 2000).

disparity or depth. Indeed, perceived depth changes were completely abolished when the binocular vergence and global disparity of a  $30^\circ \times 30^\circ$  random-dot pattern oscillated over the range of binocular fusion (Erkelens & Collewijn, 1985a, 1985b; Regan, Erkelens, & Collewijn, 1986). Perceiving motion in depth requires *relative* disparity.

How reliable is the relation between perceived depth and disparity? Two studies have shown that discriminations of depth from disparity are imprecise, and worse than those for monocular relative positions. McKee, Levi, and Bowne (1990) found that increment thresholds for depth separations were much larger than similar monocular thresholds for differences in width separation. Norman, Norman, Craft, Walton, Bartholomew, Burton, Wiesemann, and Crabtree (2008) tested depth-order discriminations for two targets at varied base disparities. Thresholds were about 22% of the base disparity.

How visible are temporal variations in disparity? Tyler (1971) found that even with a stationary reference line, disparity thresholds for motion in depth were two to four times larger than those for monocular motion at oscillation rates of 0.1–5 Hz. Monocular thresholds were below 30 arcsec for speeds of 0.5 Hz and above, but binocular disparity did not achieve hyperacuity at any oscillation rate.

How visible are gradients of disparity and depth for slanted surfaces? Stereoscopic slant perception is complicated by the fact that many factors influence relationships between surface slant and disparity and between disparity and perceived slant—especially viewing distance, but also stimulus size, vergence, cyclovergence, and direction of slant (see Howard & Rogers, 2002, chap. 21). For example, vertical surface slant (around the horizontal axis) is much more visible than horizontal slant (e.g., Gillam & Rogers, 1991; Gillam & Ryan, 1992; Howard & Kaneko, 1994; Rogers & Graham, 1983).

How does vision resolve ambiguities in the many-to-many correspondence between surface slant and disparity gradients? Some researchers have proposed that slant is perceived by averaging depth estimates from individual cues (e.g., viewing distance, horizontal shear, and vertical shear), weighting each estimate by the cue's reliability (Backus & Banks, 1999; Banks, Hooge, & Backus, 2001). Another possibility is that binocular vision is simply insensitive to spatial gradients of disparity—consistent with the inherent image ambiguity.

Indeed, stereopsis seems to add little to the limited precision of slant discriminations without disparity. Norman, Crabtree, Bartholomew, and Ferrell (2009) tested slant discriminations, with large ( $22^\circ$  diameter) random-dot stereograms at a constant viewing distance. They found that stereo slant thresholds were similar to those for texture and motion gradients without disparity. Weber fractions averaged

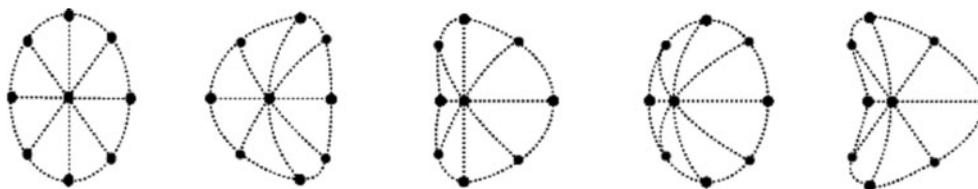
(root-mean squared [r.m.s.]) 14% for 10 observers.<sup>10</sup> Slant discriminations for these planar surfaces were more precise than those obtained when well-trained observers compared the relative slants of two regions on randomly shaped solids defined by binocular disparity, texture, and shading (Norman, Todd, Norman, Clayton, & McBride, 2006). Even with the added texture and shading cues, discrimination thresholds (at  $d' = 1.0$ ) were about  $5^\circ$ – $10^\circ$ . Stereo slant discriminations were modest in these studies but would have been worse if viewing distances were randomly varied.

*Invariance:* What is the reference frame for visual information about spatial position, binocular disparity, and motion? The issue is empirical: What groups of transformations leave perceived relative positions, depths, and motions invariant? Hyperacuties for detecting interimage disparities offer a sensitive experimental method for identifying the spatial reference frame.

If monocular images and binocular disparities were spatially defined by retinal anatomy (e.g., Marr, 1982), then stereoacuity and detection of differences in depth would vary with the fixation point, viewing distance, ocular vergence, and image motion. Experiments have clearly shown, however, that large variations in retinal disparities are (a) not perceived when applied globally over a large visual area and (b) have little to no effect on stereoscopic acuity for detecting a small local depth difference (Erkelens & Collewijn, 1985a, 1985b; Lappin & Craft, 1997, 2000; Regan et al., 1986; Steinman, Levinson, Collewijn, & van der Steen, 1985; van Ee & Erkelens, 1996; Westheimer & McKee, 1978).

Lappin and Craft (2000), for example, independently jittered each monocular image, with 10-Hz random horizontal and vertical shifts; the horizontal image shifts averaged (r.m.s.) 340 arcsec in each eye (= 455 arcsec horizontal disparity). In a stereoacuity task, observers tried to eliminate relative depth by adjusting a central vertical line to be coplanar with two parallel outer reference lines separated by  $1^\circ$ – $8^\circ$ . In a similar motion task (with no binocular disparity), observers tried to eliminate additional independent horizontal motions of the target line relative to the outer reference lines. Random image jitter had almost no effect on acuity for relative position in either task. When the two monocular images were independently jittered in the stereoacuity task, the average disparity threshold was 0.29% of the separation between the reference lines, as compared to 0.22% for stationary patterns. (With reference lines separated by  $2^\circ$ , for example, disparity thresholds were 13.2 and 12.4 arcsec for the jittered and stationary patterns,

<sup>10</sup> This experiment tested both young and older observers, but here we report performance of only the young observers. Weber fractions for slant estimations by a palm board are defined here as  $SD/M$  of the estimated slants, using the  $SD$  of slant estimations by individual observers at a given slant, data not reported in the published article.



**Fig. 7** Deformations of second-order 2-D differential structure at a given image point produced by each of four different local surface shapes when the object rotates horizontally relative to the observer. Before the rotation, the initial pattern was circular, and the local image deformations produced by the rotation are shown above. Each pattern is centered on the axis of rotation; the center is a reference point that does not move. Binocular disparities between the two monocular parts of stereoscopic images involve the same image deformations. The second-order 2-D differential structure of these patterns is invariant

under translation, 2-D rotation in the image plane, expansion or contraction, and slant of the surface relative to the image. Such image transformations produced by movements in 3-D space affect all lower-order structure, however. As can be readily seen, the five shapes, from left to right, are a plane, horizontal cylinder (*parabolic*), vertical cylinder (*parabolic*), sphere (*elliptic*), and saddle (*hyperbolic*). [The illustration is from Lappin & Craft, 2000, Fig. 3, p.14. Copyright 2000 by the American Psychological Association. Reprinted with permission.]

respectively.) Similarly, when the images were globally jittered in the motion task, acuities for relative motion averaged 0.18% of the reference line separation, as compared to a Weber fraction of 0.06% for stationary patterns.

Thus, spatial information for binocular disparity and relative motion is topological—defined by reference to the surrounding image, not by anatomical retinal positions. The simplest spatial relation involves pairs of points. Pair-wise (two-point) structure may be described by first-order spatial derivatives, Fourier power spectrum, and auto- or cross-correlation. Pair-wise spatial relations are invariant with global image translations but are perturbed by transformations such as global dilations (expansion/contraction) or 2-D rotation of the images. Lappin and Craft (1997, 2000) found that acuities for detecting local depth and relative motion were maintained when uncorrelated random dilations and image rotations were added to the random translations (Lappin & Craft, 1997, 2000). Thus, the topology of image information for stereoscopic vision and motion perception involves more than two-point relations.

Other possibilities include three-point forms, described by second-order spatial derivatives and the Fourier phase spectrum. Triangular (2-D) three-point forms are invariant under first-order image transformations—dilations and rotations—but are perturbed by changes in surface slant. Thus, three-point forms cannot provide direct information about surface shape.

As shown earlier, local surface shape is described by the second-order 2-D differential structure of surfaces and images. This second-order 2-D spatial form involves five-point relations—describing the 2-D neighborhood around any given point, as illustrated in Fig. 7.<sup>11</sup> The following

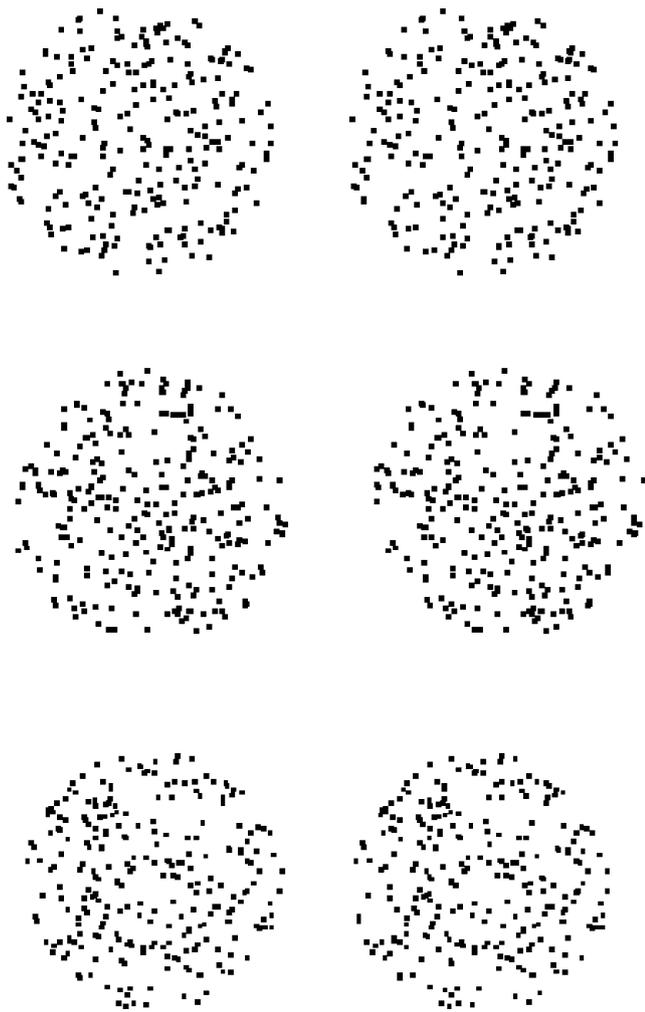
<sup>11</sup> The complexity of this elementary form is fourth-order. The “order” of complexity is one less than the number of defining points—a scalar array is zeroth-order, two-point structures are first-order, and so forth. The shape-related forms of surfaces and their images are often referred to as “second-order,” but there are two independent dimensions, involving the relative magnitudes of two second-order relations around a given point.

experiments found that human vision is directly sensitive to this higher-order spatial information about surface shape.

Perotti, Todd, Lappin, and Phillips (1998) found that precise visual resolution of local surface shape was invariant with image transformations produced by movements in 3-D space. Motion parallax fields described smooth surfaces rotating back and forth around a vertical axis. The surface shapes were randomly varied:  $z = (\kappa_1 x^2 + \kappa_2 y^2)/2$ , where  $z$  was the depth axis,  $\kappa_1$  and  $\kappa_2$  were principal curvatures, randomly varied between patterns, and  $x$  and  $y$  were the image coordinates. Schematic illustrations are shown in Fig. 8. Observers estimated the shape and curvedness of the moving surface by adjusting the curvatures ( $\kappa_{s1}$  and  $\kappa_{s2}$ ) of a stationary stereoscopic surface.

Shape indices (Eq. 6) of the observers’ adjusted stereoscopic surfaces corresponded to those of the moving standard surfaces (average  $R^2 = .995$ ), but judgments of curvedness (Eq. 7) were both imprecise (average  $R^2 = .415$ ) and inaccurate. Representative results for 1 observer are shown in Fig. 9. The precision of perceived shape and the imprecision of perceived curvedness are consistent with the image information.

Importantly, Perotti et al. (1998) tested the invariance of perceived shape by three additional conditions, with added transformations of first-order image structure—*curl* (2-D rotation), *divergence* (expansion/contraction), and *shear* (expansion and contraction in perpendicular directions, associated with slant). Curl and divergence alter the first-order spatial structure, and shear also deforms the triangular three-point forms. None of these added image transformations reduced the visual resolution of shape: Correlations between simulated and adjusted shapes in the curl, divergence, and shear conditions, respectively, averaged  $R^2 = .997$ ,  $.993$ , and  $.997$ . Visual demonstrations of this invariance appear unremarkable: The same shape simply has added rigid motions—rotating in the plane, moving closer or farther, or slanting in depth. The phenomenon is illustrated in Fig. 10.



**Fig. 8** Three stereoscopic illustrations of surface shapes used by Perotti et al. (1998) to evaluate the perception of shape from motion. (Top) Ellipsoid. (Center) Saddle (hyperbolic). (Bottom) Ridge (almost parabolic, but with slight curvature in the direction from lower right to upper left, making it hyperbolic). [From Perotti et al., 1998, Fig. 3., p.381. Copyright 1998 by the Psychonomic Society. Reprinted with permission.]

Lappin and Craft (2000) tested the invariance of stereo and motion hyperacuties for surface shape. Observers adjusted 3-D positions of target points onto planar and spherical surfaces. The patterns were perturbed by random translations, dilations, and rotations, and the surfaces were slanted in depth. The gradient for a slanted plane involves first-order spatial relations, but the doubly curved sphere involves a 2-D relationship among second-order relations.

Hyperacuties were obtained for the spherical surface as well as for the plane. Stereo thresholds (*SDs*) averaged 12 arcsec for the plane and 15 arcsec for the sphere; motion thresholds averaged 19 arcsec for the plane and 20 arcsec for the sphere. These results coincide with those of Perotti et al. (1998).

**Summary:** Visual resolution, uncertainty, and invariance of interimage disparities in stereopsis and motion perception

converge with our previous analysis of image information: Vision is directly sensitive to local surface shape, involving second-order 2-D differential structure of the image disparities. The invariance tests demonstrate that this information is not derived from lower-order image properties. Indeed, vision is not very sensitive to the lower-order properties, especially not the zero-order property of retinal position.

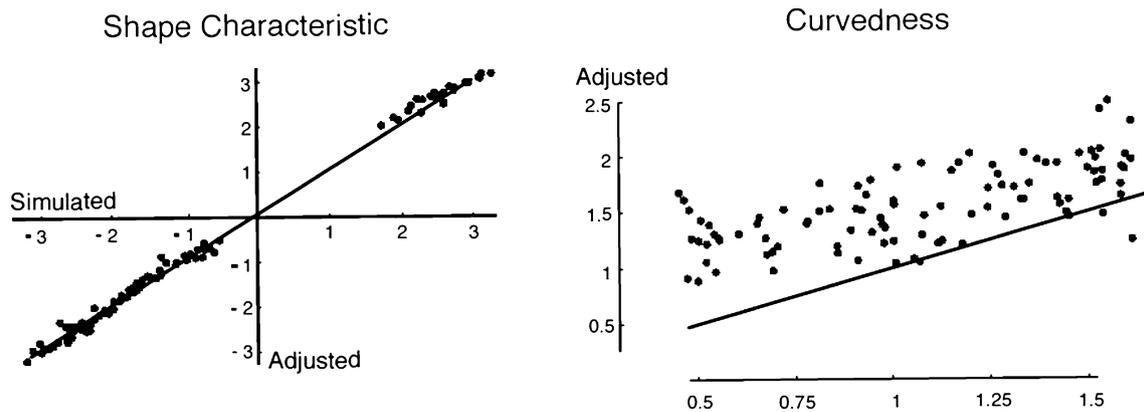
#### *Boundary contours*

Image contours correspond to surface points where the viewing direction is tangent to the surface, perpendicular to the surface normal. Thus, contour shapes are informative about surface shapes. The sign of contour curvature is the sign of the surface Gaussian curvature: Convex contours occur at convex elliptic patches (bumps, hills); concave contours, at hyperbolic (saddle-shaped) patches; inflections in contour curvature, at parabolic surface curves; and noncurved straight contours appear along edges of parabolic (cylindrical) shapes (Koenderink, 1984b, 1987; 1990, pp.431–439). A famous drawing by Picasso, in Fig. 11, is an elegant illustration. The same relationships are also shown in the left panel of Fig. 4.

Studies suggest that contours are an especially effective form of visual information about solid shape. Koenderink, van Doorn, Christou, and Lappin (1996) and Norman, Bartholomew, and Burton (2008) found that adding surface-related information such as shading, texture, and motion yielded only small improvements in shape judgments beyond those obtained with boundary contours alone. Wagemans, van Doorn, and Koenderink (2010) found that contour shape strongly influenced perception of shape from shading.

Silhouettes are devoid of all information about surface structure except for the outside boundary contours, yet they are often sufficient for perception. Attneave (1954) suggested that visual information is concentrated at contours' curvature extrema, at both convexities and concavities. Hoffman and Richards (1984) also showed that the maxima of negative curvature (concavities) of contours mark junctions of component parts. Norman, Phillips, and Ross (2001) tested these hypotheses with silhouettes of similarly shaped natural objects (potatoes), which observers tried to represent as dotted figures of 10 points. As predicted, these descriptive points were located mainly at the curvature extrema.

By definition, contour curvature involves a second-order spatial derivative. Many scientists have suggested that vision measures contour curvature by changes in (first-order) orientation (e.g., Watt & Andrews, 1982; Wilson, 1985). Koenderink and Richards (1988) and Dobbins, Zucker, and Cynader (1989), however, showed that



**Fig. 9** Representative results for 1 observer in the study of Perotti et al. (1998). The left graph shows the observer's adjusted shape characteristics as a function of the simulated (randomly generated and displayed by the computer) shape characteristic of the displayed surface pattern. The right graph shows the adjusted versus simulated

curvedness values for the same data points in the left-hand graph. Clearly, the observer accurately and reliably matched the shape but not the curvedness of these surfaces. [From Perotti et al., 1998, Fig. 4, p.382. Copyright 1998 by the Psychonomic Society. Reprinted with permission.]

second-order operators are biologically feasible and computationally efficient.

Boundary contours depend on the viewpoint; they are not fixed on the surface. When the observer or object moves, the resulting image motions are quite different from those of surface texture. Nevertheless, moving contours and silhouettes carry visible information about solid shape, as found in the experiments described below.

**Resolution:** Human vision is remarkably sensitive to contour curvature and connectedness. Experiments typically have investigated 2-D contour shapes rather than surface shapes, but the results probably generalize to surface shape.

Vision has hyperacuity for variations in contour curvature. Wilkinson, Wilson, and Habak (1998) evaluated thresholds for detecting sinusoidal modulations of the radius of circular contours. For radial frequencies (cycles per circumference) greater than 2 cycles, amplitude thresholds were a constant fraction of the radius, with Weber fractions averaging about 0.35%. Detection thresholds were invariant with contour width, and robust under reduced contrast. Modulation thresholds for circular contours were similar to those for straight lines (Tyler, 1973).

**Perceptual continuity.**—Humans are sensitive to the implied continuity of smooth contours defined by spatially separate image segments and features. Field, Hayes, and Hess (1993) demonstrated that smooth curves of separate Gabor patches could be accurately detected in dense random patterns of such patches. Perturbations of local orientations rapidly reduced detectability of the target contours. Even collinear dots will visually “pop out” of dense random-dot backgrounds (Beck, Rosenfeld, & Ivry, 1989; Uttal, 1975). Global patterns of multiple curved contours are readily visible in random-dot patterns when

neighboring pairs of dots specify smoothly varied local orientations (Glass, 1969; Wilson & Wilkinson, 1998).

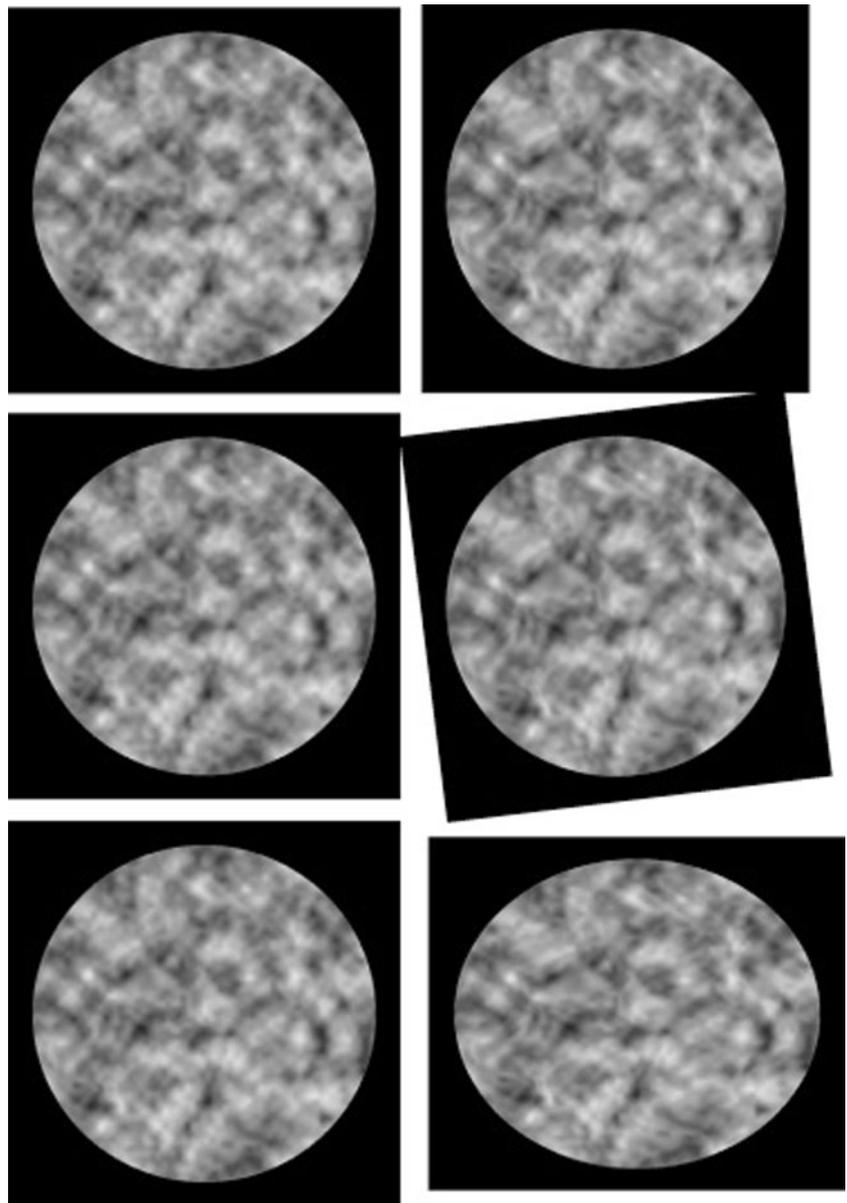
**Perceptual closure.**—Human vision is also sensitive to the global property of contour closure: Closed contours can be detected more rapidly than open contours (Elder & Zucker, 1993, 1994) and have lower contrast detection thresholds (Kovács & Julesz, 1994). The shapes of closed contours often constitute visually effective information about global shapes, especially those that are bilaterally symmetric or have distinctive appendages.

Closed contours permit a global shape description in terms of the medial axis of the outside contours. The medial axis is a type of second-order differential property—a set of points centered in a diffusion field bounded by the outside contours. Theoretical and experimental results support the visual utility of this description (Blum, 1973; Burbeck & Pizer, 1995; Elder & Zucker, 1994; Pizlo, 2008). Boundary contours and medial axis both change with the observer's viewpoint, but these deformations are informative about solid shape.

**Uncertainty.** Because contours are visually important for detecting and recognizing solid objects, contour ambiguities are important in both biology and warfare. Literatures on camouflage are relevant to the study of shape perception. A recent issue of the *Philosophical Transactions of the Royal Society B* (vol. 364, 27 February 2009) has provided helpful reviews of the biological literature. Articles by Hanlon, Chiao, Mathger, Barbosa, Buresch, and Chubb (2009), M. Stevens and Merilaita (2009a, 2009b), Tankus and Yeshurun (2009), and Troscianko, Benton, Lovell, Tolhurst, and Pizlo (2009) examined contour camouflage.

Two camouflage mechanisms that conceal boundaries are *background matching* and *disruptive coloration*. Dis-

**Fig. 10** Stereoscopic illustration of the invariance of perceived shape under added image transformations produced by 2-D rotation (“curl”) and shear (“def”). Surface shape and grayscale shading were random and mutually independent. (Top) Undistorted stereo, with right image rotated in depth around the vertical axis by about 5°. (Center) Right image rotated 7°. (Bottom) Right image horizontally expanded and vertically compressed about 7% in each axis

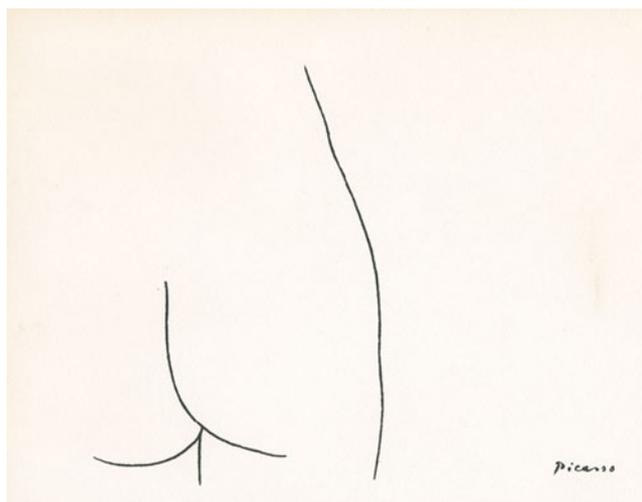


ruptive coloration adds misleading contours unrelated to an object’s shape and boundary contours. Disruptive coloration can be effective even when it increases contrast with the background (Hanlon et al., 2009).

Another potential ambiguity of contours involves the similarity between smoothly curved surface boundaries and sharp edges of planar “cutout” figures. The surface locations of image contours usually change with the observer’s viewpoint, but the edges of planar figures remain in essentially the same surface location. Thus, the image uncertainty usually disappears with changes in viewpoint. From a single viewpoint, however, the ambiguity may be important. Indeed, flat decoys resembling planes and buildings on runways when seen from above were used in WWII to misdirect or delay attacking pilots.

*Invariance:* Boundary contours and silhouettes change with the relative positions of a solid object and its image. Many experiments have found that human observers can identify and discriminate solid shapes seen only as silhouettes, even when viewpoints are varied (Newell & Findlay, 1997; Norman, Bartholomew, & Burton, 2008; Tjan, Braje, Legge, & Kersten, 1995; Wagemans et al., 2008).

The silhouettes of solid objects change when the objects rotate in depth. These contour motions are very different from those of surface texture, because they correspond to changing surface locations as the object rotates. Nevertheless, observers typically perceive the underlying object shape (Norman, Bartholomew, & Burton 2008; Norman & Raines, 2002). Norman, Lee, Phillips, Norman, Jennings, and McBride (2009) demonstrated this visual capacity by



**Fig. 11** A famous drawing by Pablo Picasso, *Fragment de corps de femme*, elegantly illustrates the correspondence between the curvature of 2-D boundary contours and 3-D surface curvature. The drawing depicts primarily elliptic (ovoid) surface regions, but also the beginning of a hyperbolic (saddle-shaped) region at the upper right, and a parabolic line at the inflection between hip and waist. Parabolic lines would also occur at the ends of the contours on the left side of the right buttock (Koenderink & van Doorn, 1982)

showing that observers accurately discriminated among similarly shaped natural objects (bell peppers) rotating in depth by seeing their moving shadows.<sup>12</sup> Observers evidently perceived invariant solids rather than changing contours as such — because their discriminations were accurate even when the shadows were projected onto curved background surfaces that altered the shapes and motions of the shadows.

### Textured images

When surfaces are covered with dense, isotropic, and homogeneous texture, image variations in the orientations, shapes, and spacing of texture depict the coordinate transformations that map surface space to image space. This image information resembles that in stereopsis and relative motion, but is also influenced by the isotropy, shapes, and contour directions of texture elements.

A visually important role of texture may be simply to mark fixed surface positions that are invariant with changes in viewpoint. Unlike many human-made objects, most natural scenes are cluttered with organic and inorganic natural materials. Textures may be inherent in the material structure of an object, or the elements may be physically

<sup>12</sup> Cast shadows and boundary contours are slightly different, though both are effective sources of information about solid shape. Both derive from space curves on the surface of solid objects, but the space curves are different because they involve tangents originating from different directions in 3-D space (Norman, Lee, et al., 2009).

distinct objects scattered over the surface; may be random or regular, contoured, polygonal, or blob-like; may be exposed by fracturing or carving solid objects that contain other shapes; and so forth. The variety of natural textures seems endless. The visual validity of assumptions about the spatial homogeneity and isotropy of surface texture seems dubious. Figure 12 shows a few examples, but these are hardly special. Importantly, perceived surface shape seems to be robust over wide variations in texture characteristics that violate theoretical assumptions such as homogeneity and isotropy (Todd & Oomes, 2002; Todd, Oomes, Koenderink, & Kappers, 2004).

Surface textures may also add image contours, and the directions of these contours may carry visual information about surface shape (e.g., Knill, 2001; Todd & Oomes, 2002; Todd & Reichel, 1990; Zaidi & Li, 2002). Contour orientation is often considered important visual information, but information about surface curvature is probably associated with contour *curvature* (a second-order space differential property, as illustrated in Figs. 7 and 11) rather than the first-order orientation as such. Todd and colleagues (Todd & Oomes, 2002; Todd & Reichel, 1990) have found that perceived shape from texture contours is quite robust.

**Resolution:** Todd et al. (2004) found that even inhomogeneous and anisotropic textures can provide reliable judgments of surface structure. Six different textures were applied to four randomly shaped solid objects. (See Fig. 13.) In one task, observers moved colored dots horizontally along a given latitudinal scan line to identify the local near and far points at that latitude. In another task, dots were equally spaced horizontally over the textured image, and observers adjusted their vertical positions on a second blank monitor to scale the perceived depths at the given positions.

Spatial resolution was not directly reported, but the reported correlations indicate that all six surface textures yielded reliable judgments of surface structure. The average judged locations of near and far points were highly correlated with the correct locations,  $r^2 = .985$ . Pair-wise correlations between different textures for the same object averaged  $r^2 = .940$ , and pair-wise correlations between observers for given stimuli averaged  $r^2 = .949$ . The average judged depths also correlated with the correct values ( $r^2 = .902$ ), but the depths were underestimated, averaging 45% of the correct values. The depth judgments of different observers ranged from 25%–62% of the correct values, although correlations between observers were high,  $r^2 = .865$ . Different textures on the same object yielded similar and correlated depth judgments ( $r^2 = .909$ ).

**Uncertainties:** A principal uncertainty of texture information is the depth scale. James Gibson (1950) and many others since have suggested that gradients (first-order spatial



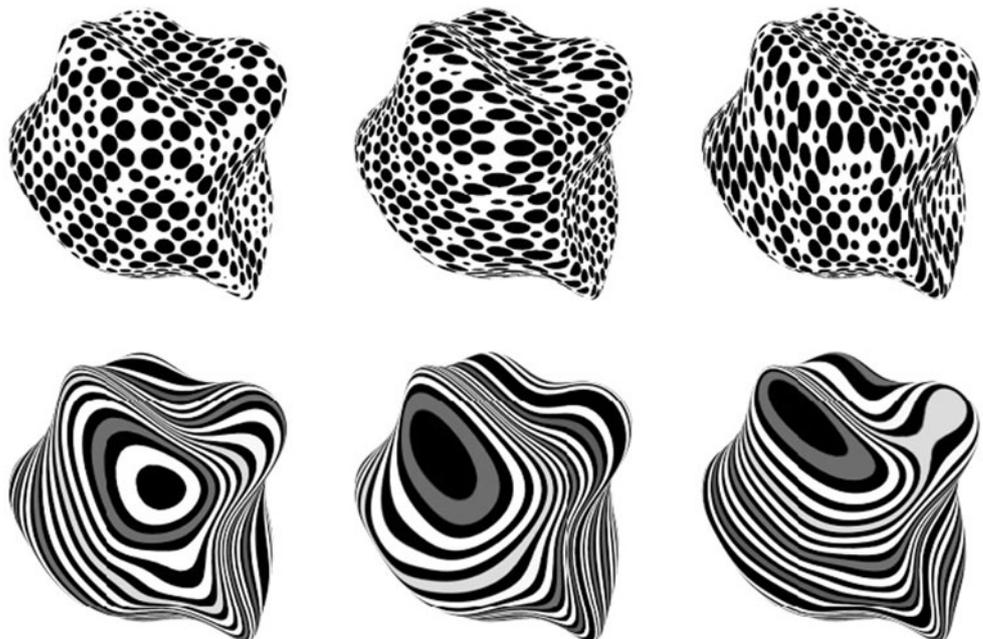
**Fig. 12** Textures provide visible markings of positions on surfaces and evidently carry information about the spatial structure of surfaces. Textures in cluttered natural scenes occur in an endless variety of configurations and scales—for instance, contours, scattered debris, volumetric textures exposed by cutting or fracturing, and pits and

holes. Uniform distributions of texture elements with the same size and shape are rare in most biological and geological structures, though more common in human-made structures. Correlations between texture patterns and surface structure are not immediately obvious in many natural scenes

derivatives) of texture density constitute information about surface slant. Slanted planes are usually clear in illustrations, but slant discriminations usually are not reliable. Norman et al. (2006), found that Weber fractions ( $SD/M$ ) for judging differences in local slant (in stereo images of textured randomly shaped solids) averaged only about 40%.

One limitation is that image texture gradients are often too small to resolve local slant. Another limitation examined by Todd and colleagues (Todd, Thaler, & Dijkstra, 2005; Todd, Thaler, Dijkstra, Koenderink, & Kappers, 2007) is that the equations relating surface slant to variations in texture density are expressed in units of visual angle. Accordingly,

**Fig. 13** Six different textures, both dotted and contoured, applied to one of the solid objects in the study of Todd et al. (2004). [From Todd et al., 2004, Fig. 1, p.41. Copyright 2004 by the Association for Psychological Science. Reprinted by Permission of SAGE Publications.]



the map from image texture to surface slant varies with the combined perspectives (focal length) of both a photograph and of an observer of the photograph; and these two perspective transformations are rarely the same.

Image size is (approximately) inversely proportional to object distance. Therefore, contrasts in image size correspond to contrasts in distance,

$$(w_N - w_F)/(w_N + w_F) \approx (d_F - d_N)/(d_F + d_N) \quad (8)$$

where  $w_N$  and  $w_F$  are the angular image widths (perpendicular to the viewing direction) of the same texture element at near and far depths, respectively, and  $d_N$  and  $d_F$  are corresponding distances from the eye (Gårding, 1992; Purdy, 1958; Todd et al., 2007). As described by Equation (8), the depth contrast for a given pair of surface elements decreases as viewing distance increases. Therefore, uncertainty about slant increases with uncertainty about viewing distance.

In recent experiments, Todd et al. (2007) found that observers' judgments of slant in computer-generated images of hyperbolic cylinders were accurately predicted by Equation (8). Slant judgments were reliable—with an average correlation of .95 between judgments of the same stimuli in different experimental sessions—but judged slants deviated systematically from the correct values. Slant judgments varied with perspective and with convexity/concavity; and depths of concave surfaces were greatly underestimated.

*Invariance:* Insofar as image information for perceiving shape from texture may involve the deformations of second-order structure described in Fig. 7, this information should be invariant under the image transformations produced by perspective changes associated with distance, direction, and focal length. The studies of Todd and colleagues (Todd & Oomes, 2002; Todd et al., 2004; Todd & Reichel, 1990; Todd et al., 2007) have suggested that perceived shape from texture may satisfy invariance under these image transformations, but such invariance has not been tested directly.

### Image shading

Surface curvature causes variations in image shading, but two other global factors also influence shading: the irradiating field of light and the scattering of light by the surface material. Unlike shape information from stereopsis, motion parallax, or boundary contours, shading information is global, not local! Koenderink and van Doorn (2004) have reviewed the fundamentals of image shading. Here, we highlight selected basic aspects.

Irradiance converges at a given surface location from a hemisphere of directions. This irradiance field is composed

of (a) direct illumination, (b) indirect reflection from other surfaces, and (c) shadows and partial occlusion. Direct illumination can be approximately unidirectional (e.g., direct sunlight), from multiple local or extended sources (e.g., rooms with multiple lamps), from a diffuse hemisphere (e.g., overcast sky), or countless combinations. Additionally, the indirect contributions of reflections, shadows, and partial occlusion by neighboring surfaces may be as great as the direct illumination. The irradiation at a location in a natural scene is a flow field structured in part by the surrounding scene. Image shading is an ecological phenomenon (Gibson, 1950, 1979; Koenderink et al., 2007; van Doorn, Koenderink, & Wagemans, 2011).

Reflection and scattering from a given surface material are described by its bidirectional reflectance distribution function (BRDF). The BRDF is the ratio of radiance scattered in a given direction relative to the incident irradiance from a different given direction. Thus, for a given surface location, the BRDF ratio (output radiance/input irradiance) depends on four independent parameters—the azimuth and elevation angles of both the incident and reflected rays. The directions of the incident and reflected rays correspond to points on a hemisphere in which the central pole is the surface normal, and BRDF ratios are defined on the four-dimensional product of two such hemispheres of directions. The BRDF generally also depends on wavelength, and it is an aspect of color vision, but we ignore that aspect here. The BRDF is the visual signature of a surface material (Koenderink & van Doorn, 1998; Koenderink, van Doorn, Dana, & Nayar, 1999; Oren & Nayar, 1995), but the complexity of the BRDF makes it difficult to measure.

An idealized model known as a *Lambertian* surface scatters light uniformly in all directions, independent of the directions of reflection and incidence. The reflected radiance (photons/time/area) is then proportional to the incident irradiance, and the BRDF is a constant. In this special case, image intensity decreases with the angle between the surface normal and the direction of irradiation, providing information about variations in surface orientation.<sup>13</sup> This model is conceptually simple but physically unusual.

Surface scattering increases with the microscopic roughness of a surface (Tadin, Haglund, Lappin, & Peters, 2001). Very smooth surfaces (metals, glass, water) are shiny, reflecting light like a mirror in a tightly constrained bundle of rays. Such shiny surfaces are called *specular*. Surfaces are sometimes modeled in computer graphics and psychophysical experiments as the sum of two components,

<sup>13</sup> The surface area irradiated by a source in a given direction increases proportionally with the cosine of the angle of incidence. Thus, the radiant intensity reflected from a given surface location decreases inversely with the angle of incidence.

Lambertian plus specular, but this is an oversimplification for most surfaces. Scattering usually is centered on the specular direction, with greater dispersion for rougher surfaces.

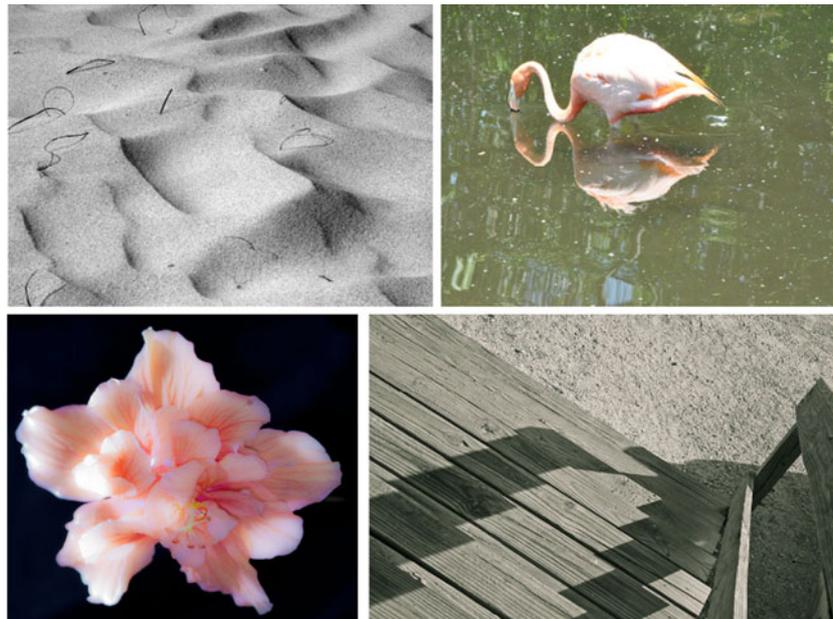
Specular surfaces produce *highlights*—images of incident light from an angle of incidence equal to the angle of reflection. Thus, highlights carry information about surface orientation. Highlights aid perception of surface shape (Norman, Todd, & Orban, 2004).

Variations in shading carry image information about microscopic roughness, macroscopic texture, surface color, and boundary contours. Figure 14 offers a few illustrations. Shading is not an independent physical cue. Roughness, texture, and boundary contours may be considered aspects of image shading information. Experimental separation of shading from other surface properties may be misleading.

*Resolution:* The current experimental literature does not establish the visual resolution of surface shape from shading. Koenderink and van Doorn (2004) concluded, “There is remarkably little psychophysical material on shape from shading that might be considered in any way definitive” (p.1100). “Perhaps the majority of the literature is irrelevant due to incomplete description of the stimuli, ... extreme stimulus reduction, ... or invalid paradigms due to incomplete understanding of the relevant geometry and photometry” (pp.1101–1102).

When variations in local shape are experimentally isolated from global information about the irradiance distribution and surrounding surface structure, local shape is very poorly resolved (Erens, Kappers, & Koenderink, 1993a, 1993b). In the Erens et al. (1993b) study, observers’ categorical shape judgments (eight regions of the shape index—Eq. 6, Fig. 5) were essentially random. When a cast shadow revealed the illumination direction, then concave and convex shapes were discriminated, but elliptic and hyperbolic shapes were still confused.

Perceiving shape from shading involves simultaneous perception of both the surface and the surrounding light field. A recent study by Koenderink et al. (2007) found that vision is quite sensitive to the structure of light fields in natural scenes. Observers adjusted the direction, diffuseness, and intensity of illumination of a gauge object (Lambertian sphere) at various positions in a multiobject scene. They easily and accurately matched the illumination of the gauge object to the local physical parameters of the light field. Correlations ( $r^2$ ) between the physically correct and adjusted shading values averaged (over conditions and observers) .77 for the direction and diffuseness parameters, and .42 for the intensity. Quantitative details suggested that shaded images of natural scenes support perception of local surface shape, but shape discriminations were not evaluated in this experiment.



**Fig. 14** Images are composed of light reflected and scattered from surfaces. Image shading depends on the surface orientation but also on many other factors, including the spatial distribution of illumination, the direction of view, shadows, reflections from other surfaces, occluding transparent materials, and the reflectance characteristics of the surface. Sand, water, wood, flowers, and feathers have very

different reflectance characteristics. Roughly textured sand and wood scatter light broadly; the smooth surface of water reflects light in a narrow distribution of directions; and scattering by flowers and feathers is intermediate between these extremes. The flower at the lower left is illuminated from behind, thus appearing luminous. [The two bottom photos were contributed by Dominic Ali.]

*Uncertainties:* Varied combinations of surface shape, BRDF, and illumination may produce the same patterns of image shading. Like most forms of image information about surface structure, both the absolute distance and relative depth variations across surface regions are ambiguous in shaded images.

Belhumeur et al. (1999) presented a clear analysis and demonstration of this “bas-relief ambiguity” of depth from shading. This ambiguity includes convexity versus concavity and, in certain cases, even connections in depth (van Doorn et al., 2011).

Perceived convexities and concavities of surface regions are known to vary with perceived directions of illumination (e.g., Kleffner & Ramachandran, 1992; Ramachandran, 1988; van Doorn et al., 2011). A popular idea is that perceived convexity versus concavity involves an automatic visual “assumption” or Bayesian “prior” that illumination is from above. The important point, however, is that the image information is inherently ambiguous. Moreover, experiments have found significant variability between individuals, within individuals, strong effects of the outer contour shape, and both interactions and inconsistencies among spatially separate surface regions (Kleffner & Ramachandran, 1992; van Doorn et al., 2011; Wagemans, van Doorn, & Koenderink, 2010).

*Invariance:* In general, perceived surface shape is not invariant with changes in shading (Koenderink & van Doorn, 2004). Changing the illumination of a given surface may produce locally varied changes in relative depth and in convexity versus concavity. Often, the changes are subtle, with the topographical complexity and global shape appearing similar after a change of illumination, but changes in local details may become visible under closer scrutiny. Careful control of illumination is important to professional photographers, and commercial success in the cosmetics industry involves controlling skin reflectance to produce perceived changes in facial structure.

Locations of highlights vary with viewpoint. When a shiny surface is viewed stereoscopically, the highlight can sometimes be seen in depth off the surface. This stereo effect can alter the perceived shape of a shiny surface (Murry, van Mierlo, Fleming, & Welchman, 2011; Nefs, 2008), though vision is usually robust to such effects. Surface texture, scattering, and diffuse illumination tend to counteract the potential uncertainties caused by the viewpoint-dependent effects of highlights (Nefs, 2008; Todd, Norman, Koenderink, & Kappers, 1997). Norman, Todd, and Phillips (1995) found that image transformations produced by binocular vision and observer/object motions aided in recovering invariant surface structure from shaded images. Perceived shapes of shaded surfaces are usually stable under relative movements of observer and object and

under changes in illumination (Koenderink & van Doorn, 1980), but this tendency is not a generalized invariance.

### *Combining multiple forms of image information*

Information about surface structure is carried by multiple image variables—binocular disparity, motion parallax, boundary contours, texture, shading, and so forth. These variables carry amounts of information that vary both within and between scenes, and between sensory channels that detect this information. How, then, does vision combine multiple sources of information? This is a basic issue in sensory and perceptual research.

To resolve this issue, one must first identify the information to be combined, but this first step has received insufficient attention. Different assumptions about the visual input and perceptual output yield differing conceptions of information integration.

Two contrasting approaches are *statistical* and *intrinsic constraint* models. The extensive theoretical and experimental literature on this topic is beyond the scope of this article, but the underlying conceptions of image information are quite relevant.

The statistical approach includes a variety of models that aim to identify quantitative rules for combining multiple forms of sensory information to infer surface structure—typically slant or depth. These rules usually use Bayesian statistical analyses of correlations between image cues and environmental properties. A simple but frequently used version assumes that independent estimates of relative depth are derived from separate sensory channels that detect different image cues (e.g., motion parallax, binocular disparity, or texture foreshortening). Typically, these cues involve zero- or first-order image properties, which are used to estimate surface depth or slant. In a simple linear version, evidence from multiple cues is linearly combined, weighting each cue by its statistical reliability and by other nonvisual information. Because these low-order image and surface properties are not reliably correlated across variations in viewpoint, estimates are improved by (a) combining evidence from multiple cues, (b) statistics of environmental and image variables, and (c) adding evidence from prior probabilities of certain environmental properties. Among many relevant studies are Landy, Maloney, Johnston, and Young (1995); Jacobs (1999); Mamassian and Landy (2001); Ernst and Banks (2002); Hillis, Watt, Landy, and Banks (2004); van Ee, Adams, and Mamassian (2003); Knill (2003); Knill and Saunders (2003); and Adams and Mamassian (2004).

An intrinsic constraint model developed by Domini and colleagues (e.g., Caudek, Fantoni, & Domini, 2011; Di Luca, Domini, & Caudek, 2007; Domini & Caudek, 2009,

2010; Domini, Caudek, & Tassinari, 2006; Vuong, Domini, & Caudek, 2006) is based on a different representation of image information and perception. Here, multiple image variables are constrained by the same underlying surface structure and converge on the same ordinal relations in depth, despite possible differences in resolution, uncertainty, or sensory modality. Image information in this model is coherent (correlated) among neighboring image points and surface points, and multiple image cues refer to the same ordinal relations among points on a given surface.

The rationale and evidence reviewed here are consistent with the intrinsic constraint model but not with the statistical models. The statistical models are limited by the image and surface information they use, not by the statistical processes that combine this information. Information may indeed be statistical, but the fruitfulness of statistical analysis depends on the variables to which it is applied. The lower-order properties used by most current statistical models are not invariant with changes in viewpoint, and statistical analyses will not overcome this limitation.

The intrinsic constraint model derives its power from information about surface shape—where multiple image cues describe the same structure, independent of the scale of resolution. Linear integration of these multiple structures converges on a common structure (Koenderink & van Doorn, 1997).

*Resolution:* Suppose that information from an unreliable depth cue is added to that from a cue with greater reliability and greater visual resolution. To what extent does the added information increase the perceptual resolution of relative depths and surface structure?

According to many statistical models (e.g., the “modified weak fusion” model of Landy et al., 1995), depth maps are estimated independently from each cue. When depth estimates from the individual cues are combined as a weighted average, the variance of the combined-cue estimate will be slightly lower even when the second cue has low resolution (Jacobs, 1999).

In contrast, the intrinsic constraint model can predict significantly better depth discrimination with added low-resolution information, because two different variables can provide complementary constraints on the surface shape (Vuong et al., 2006). Surface shape may be described independently of the image resolution, structurally compatible across resolution scales (Koenderink, 1990).

Vuong et al. (2006) tested these contrasting predictions by comparing the precision with which a stereoscopic probe dot could be adjusted in depth along the surface normal to lie on a curved surface specified by either disparity (of a sparse stereo-dot array), monocular shading (Lambertian), or both. In one experiment, the shading was identical in both eyes, providing no information for the stereo adjust-

ment task; in another experiment, the boundary contours of the shaded image differed slightly between the left and right images, providing a weak stereo cue. In both experiments, the precision (*SD*) of the depth adjustments averaged 15 arcsec in the disparity-only condition and improved to 11 arcsec in the combined condition, even when the monocular shading provided no stereoscopic information at all. When the shaded contours differed slightly between the two stereo images, the precision of depth adjustments in the shading-only condition averaged 35 arcsec, and the statistically predicted performance in the combined condition (according to the “modified weak fusion” model) was 13 arcsec. Adjustments in the combined-cue condition by all 4 observers were about 20% more precise than predicted by the statistical model.

Vuong et al.’s results indicate that surface information from low-resolution shading and high-resolution disparities was structurally compatible. Whereas the stereo-dot array was sparse, the shading was continuous and aided visual surface interpolation between dots.

*Invariance:* The strength of the intrinsic constraint model reflects the use of surface structure that is isomorphic across image properties (e.g., disparity, shading, motion, etc.) and invariant with viewing conditions such as viewpoint, illumination, and resolution.

## Summary and conclusion

This article began with the idea that Fechner’s insight about relations between matter and mind is relevant to the fundamental problem of vision—concerning how retinal images yield visually perceived objects. Fechner’s insight was that relations between matter and mind involve corresponding structures of variations in the material and mental worlds. Fechner’s idea is one version of a general concept of information, involving the representation of a relational structure in one domain by that in another. Thus, a basic goal of vision science is to identify the elementary spatial structure that is shared in common by environmental objects, optical images, and perceptual discriminations.

Three criteria for evaluating representation of structural relations are (a) resolution, (b) uncertainty, and (c) invariance. For spatial vision, invariance under changes in observational parameters such as viewpoint, focal length, illumination, and context is especially relevant.

Applying these criteria to theoretical and experimental evidence about spatial vision reveals that the structural correspondence between environmental objects, optical images, and perceptual discriminations is the differential structure associated with *local surface shape*. Other spatial

properties involving relative depth and surface slant do not satisfy these criteria and, therefore, do not constitute reliable information for seeing environmental objects.

The general conclusion is that a modern version of Fechner's idea about the relation between matter and mind is a key to understanding spatial vision.

**Author note** The authors are grateful to Jan Koenderink and James Todd for helpful discussions and insightful suggestions on several aspects of this article; to Ragnar Steingrímsson and Joshua Solomon for helpful comments on an earlier version; and for the extensive and insightful recommendations of two anonymous reviewers. We are also grateful to Douglas Morse for his help with German-to-English translation of extensive sections of the book on *Foundations and Functions of Information Theory* by Meyer-Eppler, and to Dominic Ali for his expert help in supplying and editing several photographs.

## References

- Adams, W. J., & Mamassian, P. (2004). Bayesian combination of ambiguous shape cues. *Journal of Vision*, *7*, 921–929. doi:10.1167/4.10.7.4(10).
- Anderson, B. L., & Winawer, J. (2008). Layered image representations and the computation of surface lightness. *Journal of Vision*, *18*, 1–22. doi:10.1167/8.7.18.8(7).
- Ashby, W. R. (1963). *An introduction to cybernetics*. New York: Wiley.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*, 183–193. doi:10.1037/h0054663
- Backus, B. T., & Banks, M. S. (1999). Estimator reliability and distance scaling in stereoscopic slant perception. *Perception*, *28*, 217–242. doi:10.1068/p2753
- Banks, M. S., Hooge, I. T. C., & Backus, B. T. (2001). Perceiving slant about a horizontal axis from stereopsis. *Journal of Vision*, *1*, 55–79. doi:10.1167/1.2.1.1(2).
- Beck, J., Rosenfeld, A., & Ivry, R. (1989). Line segregation. *Spatial Vision*, *4*, 75–101.
- Belhumeur, P. N., Kriegman, D. J., & Yuille, A. L. (1999). The bas-relief ambiguity. *International Journal of Computer Vision*, *35*, 33–44.
- Blake, R. (1993). Cats perceive biological motion. *Psychological Science*, *4*, 54–57. doi:10.1111/j.1467-9280.1993.tb00557.x
- Blum, H. (1973). Biological shape and visual science (Part I). *Journal of Theoretical Biology*, *38*, 205–287.
- Braunstein, M. L. (1962). Depth perception in rotating dot patterns: Effects of numerosity and perspective. *Journal of Experimental Psychology*, *64*, 415–420. doi:10.1037/h0048140
- Burbeck, C. A., & Pizer, S. M. (1995). Object recognition by cores: Identifying and representing primitive spatial regions. *Vision Research*, *35*, 1917–1930. doi:10.1016/0042-6989(94)00286-U
- Caudek, C., Fantoni, C., & Domini, F. (2011). Bayesian modeling of perceived surface slant from actively-generated and passively-observed optic flow. *PLoS One*, *6*, e18731.
- Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision Research*, *31*, 2195–2207. doi:10.1016/0042-6989(91)90172-2
- Cutting, J. E. (1987). Rigidity in cinema seen from front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 323–334.
- Descartes, R. (1886). *La géométrie*. Paris: A. Hermann. (Original work published 1637)
- De Valois, K. K., Lakshminarayanan, V., Nygaard, R., Schluskel, S., & Sladky, J. (1990). Discrimination of relative spatial position. *Vision Research*, *30*, 1649–1660. doi:10.1016/0042-6989(90)90150-J
- Di Luca, M., Domini, F., & Caudek, C. (2007). The relation between disparity and velocity signals of rigidly moving objects constrains depth order perception. *Vision Research*, *47*, 1335–1349. doi:10.1016/j.visres.2006.10.029
- Dobbins, A., Zucker, S. W., & Cynader, M. S. (1989). Endstopping and curvature. *Vision Research*, *29*, 1371–1387.
- Domini, F., & Caudek, C. (2009). The intrinsic constraint model and Fechnerian sensory scaling. *Journal of Vision*, *25*, 1–15. doi:10.1167/9.2.25.9(2).
- Domini, F., & Caudek, C. (2010). Matching perceived depth from disparity and velocity: Modeling and psychophysics. *Acta Psychologica*, *133*, 81–89.
- Domini, F., Caudek, C., & Tassinari, H. (2006). Stereo and motion information are not independently processed by the visual system. *Vision Research*, *46*, 1707–1723.
- Elder, J., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research*, *33*, 981–991. doi:10.1016/0042-6989(93)90080-G
- Elder, J., & Zucker, S. (1994). A measure of closure. *Vision Research*, *34*, 3361–3369. doi:10.1016/0042-6989(94)90070-1
- Erens, R. G. F., Kappers, A. M. L., & Koenderink, J. J. (1993a). Estimating local shape from shading in the presence of global shading. *Perception & Psychophysics*, *54*, 334–342.
- Erens, R. G. F., Kappers, A. M. L., & Koenderink, J. J. (1993b). Perception of local shape from shading. *Perception & Psychophysics*, *54*, 145–156. doi:10.3758/BF03211750
- Erkelens, C. J., & Collewijn, H. (1985a). Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. *Vision Research*, *25*, 1689–1700. doi:10.1016/0042-6989(85)90141-5
- Erkelens, C. J., & Collewijn, H. (1985b). Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Research*, *25*, 583–588. doi:10.1016/0042-6989(85)90164-6
- Ernst, M. O., & Banks, M. S. (2002). Human integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.
- Fechner, G. T. (1860). *Elemente der Psychophysik vol. 2*. Leipzig: Breitkopf & Hartel.
- Fechner, G. T. (1965). *Elemente der Psychophysik*. In R. J. Herrnstein & E. G. Boring (Eds.), *A source book in the history of psychology*. Cambridge, MA: Harvard University Press (Original work published 1860).
- Fechner, G. T. (1966). *Elements of psychophysics, Vol. 1 (H.E. Adler; Trans.)*. New York: Holt, Rinehart & Winston (Original work published 1860).
- Fechner, G. T. (2004). *Elemente der Psychophysik, Vol. 2 (C. Kloth, Trans.)*. In M. Heidelberger (Ed.), *Nature from within: Gustav Theodor Fechner and his psychophysical worldview*. Pittsburgh, PA: University of Pittsburgh Press (Original work published 1860).
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local “association field”. *Vision Research*, *33*, 173–193.
- Florack, L. (1997). *Image structure*. Dordrecht, The Netherlands: Kluwer.
- Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science*, *218*, 486–487.
- Gårding, J. (1992). Shape from texture for smooth curved surfaces in perspective projection. *Journal of Mathematical Imaging and Vision*, *2*, 327–350.
- Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.
- Garner, W. R., Hake, H. W., & Eriksen, C. W. (1956). Operationism and the concept of perception. *Psychological Review*, *63*, 149–159. doi:10.1037/h0042992

- Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gillam, B., & Rogers, B. (1991). Orientation disparity, deformation, and stereoscopic slant perception. *Perception, 20*, 441–448.
- Gillam, B., & Ryan, C. (1992). Perspective, orientation disparity, and anisotropy in stereoscopic slant perception. *Perception, 21*, 427–439.
- Glass, L. (1969). Moiré effect from random dots. *Nature, 223*, 578–580. doi:10.1038/223578a0
- Gray, A. (1993). *Modern differential geometry of curves and surfaces*. Boca Raton, FL: CRC Press.
- Hanlon, R. T., Chiao, C.-C., Mathger, L. M., Barbosa, A., Buresch, K. C., & Chubb, C. (2009). Cephalopod dynamic camouflage: Bridging the continuum between background matching and disruptive coloration. *Philosophical Transactions of the Royal Society B, 364*, 429–437.
- Heidelberger, M. (2004). *Nature from within: Gustav Theodor Fechner and his psychophysical worldview*. Pittsburgh: University of Pittsburgh Press.
- Hess, R. F. (1982). Developmental sensory impairment: amblyopia or tarachopia? *Human Neurobiology, 1*, 17–29.
- Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision, 1*, 967–992. doi:10.1167/4.12.1. 4(12).
- Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition, 18*, 65–96.
- Howard, I. P., & Kaneko, H. (1994). Relative shear disparities and the perception of surface inclination. *Vision Research, 34*, 2505–2517. doi:10.1016/0042-6989(94)90237-2
- Howard, I. P., & Rogers, B. J. (2002). *Seeing in depth* (Vol. 2: Depth perception). Thornhill, Ontario, Canada: I. Porteous.
- Ittelson, W. H. (1952). *The Ames demonstrations in perception*. Princeton, NJ: Princeton University Press.
- Jacobs, R. (1999). Optimal integration of texture and motion cues to depth. *Vision Research, 39*, 3621–3629.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14*, 201–211.
- Julesz, B., Pappathomas, T. V., & Phillips, F. (2006). *Foundations of cyclopean perception*. Cambridge, MA: MIT Press.
- Kleffner, D. A., & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics, 52*, 18–36. doi:10.3758/BF03206757
- Knill, D. C. (2001). Contour into texture: information content of surface contours and texture flow. *Journal of the Optical Society of America, A, 18*, 12–35.
- Knill, D. C. (2003). Mixture models and the probabilistic structure of depth cues. *Vision Research, 43*, 831–854.
- Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43*, 2539–2558.
- Koenderink, J. J. (1984a). The structure of images. *Biological Cybernetics, 50*, 363–370.
- Koenderink, J. J. (1984b). What does the occluding contour tell us about solid shape? *Perception, 13*, 321–330.
- Koenderink, J. J. (1987). Internal representation of solid shape based on the topological properties of the apparent contour. In W. Richards & S. Ullman (Eds.), *Image understanding 1985–1986* (pp. 257–285). Norwood, NJ: Ablex.
- Koenderink, J. J. (1990). *Solid shape*. Cambridge, MA: MIT Press.
- Koenderink, J. J. (2001). Multiple visual worlds. *Perception, 30*, 1–7.
- Koenderink, J. J., Pont, S. C., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2007). The visual light field. *Perception, 36*, 1595–1610.
- Koenderink, J. J., & Richards, W. (1988). Two-dimensional curvature operators. *Journal of the Optical Society of America, A, 5*, 1136–1141.
- Koenderink, J. J., & van Doorn, A. J. (1975). Invariant properties of motion parallax due to the movement of rigid bodies relative to the observer. *Optica Acta, 22*, 773–791.
- Koenderink, J. J., & van Doorn, A. J. (1976a). Geometry of binocular vision and a model for stereopsis. *Biological Cybernetics, 21*, 29–35.
- Koenderink, J. J., & van Doorn, A. J. (1976b). Local structure of movement parallax of the plane. *Journal of the Optical Society of America, 66*, 717–723.
- Koenderink, J. J., & van Doorn, A. J. (1976c). The singularities of the visual mapping. *Biological Cybernetics, 24*, 51–59.
- Koenderink, J. J., & van Doorn, A. J. (1980). Photometric invariants related to solid shape. *Optica Acta, 27*, 981–986.
- Koenderink, J. J., & van Doorn, A. J. (1982). The shape of smooth objects and the way contours end. *Perception, 11*, 129–137.
- Koenderink, J. J., & van Doorn, A. J. (1992a). Generic neighborhood operators. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 14*, 597–605.
- Koenderink, J. J., & van Doorn, A. J. (1992b). Second-order optic flow. *Journal of the Optical Society of America, A, 9*, 530–538.
- Koenderink, J. J., & van Doorn, A. J. (1992c). Surface shape and curvature scales. *Image and Vision Computing, 10*, 557–564.
- Koenderink, J. J., & van Doorn, A. J. (1997). The generic bilinear calibration-estimation problem. *International Journal of Computer Vision, 23*, 217–234.
- Koenderink, J. J., & van Doorn, A. J. (1998). Phenomenological description of bidirectional surface reflection. *Journal of the Optical Society of America, A, 15*, 2903–2912.
- Koenderink, J. J., & van Doorn, A. J. (2004). Shape and shading. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1090–1105). Cambridge, MA: MIT Press.
- Koenderink, J. J., van Doorn, A. J., Christou, C., & Lappin, J. S. (1996). Shape constancy in pictorial relief. *Perception, 25*, 155–164.
- Koenderink, J. J., van Doorn, A. J., Dana, K. J., & Nayar, S. (1999). Bidirectional reflectance distribution function of thoroughly pitted surfaces. *International Journal of Computer Vision, 31*, 129–144.
- Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1996b). Pictorial surface attitude and local depth comparisons. *Perception & Psychophysics, 58*, 163–173.
- 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.
- Koenderink, J. J., van Doorn, A. J., & Todd, J. T. (2009). Wide distribution of external local sign in the normal population. *Psychological Research, 73*, 14–22.
- Kovács, E., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature, 370*, 644–646.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement: Vol. 1. Additive and polynomial representations*. New York: Academic Press.
- Laming, S. (2010). Statistical information and uncertainty: A critique of applications in experimental psychology. *Entropy, 12*, 720–771.
- Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35*, 389–412.
- Lappin, J. S., & Bell, H. H. (1976). The detection of coherence in moving visual patterns. *Vision Research, 16*, 161–168.
- Lappin, J. S., & Craft, W. D. (1997). Definition and detection of binocular disparity. *Vision Research, 37*, 2953–2974.
- Lappin, J. S., & Craft, W. D. (2000). Foundations of spatial vision: From retinal images to perceived shapes. *Psychological Review, 107*, 6–38.

- Lappin, J. S., Shelton, A. L., & Rieser, J. J. (2006). Environmental context influences visually perceived distance. *Perception & Psychophysics*, *68*, 571–581.
- Lappin, J. S., Tadin, D., Nyquist, J. B., & Corn, A. L. (2009). Spatial and temporal limits of motion perception across variations in speed, eccentricity, and low vision. *Journal of Vision*, *30*, 1–14. doi:10.1167/9.1.30. 9(1).
- Lederman, L. M., & Hill, C. T. (2004). *Symmetry, and the beautiful universe*. Amherst, NY: Prometheus.
- Luce, R. D. (2004). Symmetric and asymmetric matching of joint presentations. *Psychological Review*, *111*, 446–454.
- Luce, R. D., Krantz, D. H., Suppes, P., & Tversky, A. (1990). *Foundations of measurement: Vol 3, Representation, axiomatization, and invariance*. New York: Academic Press.
- Mamassian, P., & Landy, M. S. (2001). Interaction of visual prior constraints. *Vision Research*, *41*, 2653–2668.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman.
- McKee, S. P., Levi, D. M., & Bowne, S. F. (1990). The imprecision of stereopsis. *Vision Research*, *30*, 1763–1779.
- Meyer-Eppler, W. (1969). *Grundlagen und Anwendungen der Informationstheorie*. Berlin: Springer.
- Muryy, A. A., van Mierlo, C. M., Fleming, R. W., & Welchman, A. E. (2011). The perception of 3D shape from binocular views of specular objects [Abstract 63.437]. *VSS 2011 Abstracts, Vision Sciences Society Annual Meeting*, 330–331.
- Nefs, H. T. (2008). Three-dimensional object shape from shading and contour disparities. *Journal of Vision*, *11*, 1–16. doi:10.1167/8.11.11. 8(11).
- Newell, F. N., & Findlay, J. M. (1997). The effect of depth rotation on object identification. *Perception*, *26*, 1231–1257.
- Norman, J. F., Bartholomew, A. N., & Burton, C. L. (2008). Aging preserves the ability to perceive 3D object shape from static but not deforming boundary contours. *Acta Psychologica*, *129*, 198–207.
- Norman, J. F., Crabtree, C. E., Bartholomew, A. N., & Ferrell, E. L. (2009). Aging and the perception of slant from optical texture, motion parallax, and binocular disparity. *Attention, Perception, & Psychophysics*, *71*, 116–130.
- Norman, J. F., Crabtree, C. E., Clayton, A. M., & Norman, H. F. (2005). The perception of distances and spatial relationships in natural outdoor environments. *Perception*, *34*, 1315–1324.
- Norman, J. F., Lee, Y., Phillips, F., Norman, H. F., Jennings, L. R., & McBride, T. R. (2009). The perception of 3-D shape from shadows cast onto curved surfaces. *Acta Psychologica*, *131*, 1–11.
- Norman, J. F., Norman, H. F., Craft, A. E., Walton, C. L., Bartholomew, A. N., Burton, C. L., Wiesemann, E. Y. & Crabtree, C. E. (2008). Stereopsis and aging. *Vision Research*, *48*, 2456–2465.
- Norman, J. F., Norman, H. F., Lee, Y., Stockton, D., & Lappin, J. S. (2004). The visual perception of length along intrinsically curved surfaces. *Perception & Psychophysics*, *66*, 77–88.
- Norman, J. F., Phillips, F., & Ross, H. E. (2001). Information concentration along the boundary contours of naturally shaped solid objects. *Perception*, *30*, 1285–1294.
- Norman, J. F., & Raines, S. R. (2002). The perception and discrimination of local 3-D surface structure from deforming and disparate boundary contours. *Perception & Psychophysics*, *64*, 1145–1159.
- Norman, J. F., & Todd, J. T. (1996). The discriminability of local surface structure. *Perception*, *25*, 381–398.
- Norman, J. F., & Todd, J. T. (1998). Stereoscopic discrimination of interval and ordinal depth relations on smooth surfaces and in empty space. *Perception*, *27*, 257–272.
- Norman, J. F., Todd, J. T., Norman, H. F., Clayton, A. M., & McBride, T. R. (2006). Visual discrimination of local surface structure: Slant, tilt, and curvedness. *Vision Research*, *46*, 1057–1069.
- Norman, J. F., Todd, J. T., & Orban, G. A. (2004). Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychological Science*, *15*, 565–570.
- Norman, J. F., Todd, J. T., Perotti, V. J., & Tittle, J. S. (1996). The visual perception of three-dimensional length. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 173–186.
- Norman, J. F., Todd, J. T., & Phillips, F. (1995). The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics*, *57*, 629–636.
- Oren, M., & Nayar, S. K. (1995). Visual appearance of thoroughly matte surfaces. *Science*, *267*, 1153–1156.
- Perotti, V. J., Todd, J. T., Lappin, J. S., & Phillips, F. (1998). The perception of surface curvature from optical motion. *Perception & Psychophysics*, *60*, 377–388.
- Phillips, F., & Todd, J. T. (1996). Perception of local three-dimensional shape. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 930–944.
- Phillips, F., Todd, J. T., Koenderink, J. J., & Kappers, A. M. L. (2003). Perceptual representation of visible surfaces. *Perception & Psychophysics*, *65*, 747–762.
- Pizlo, Z. (2008). *3D shape, its unique place in visual perception*. Cambridge, MA: MIT Press.
- Pont, S. (2011). An ecologically valid description of the light field [Abstract 26.302]. *VSS 2011 Abstracts, Vision Sciences Society Annual Meeting*, p.84.
- Purdy, W. P. (1958). The hypothesis of psychophysical correspondence in space perception. *Dissertation Abstracts*, *42*, 1454. (UMI No. 58–5594).
- Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, *331*, 163–166. doi:10.1038/331163a0
- Regan, D., Erkelens, C. J., & Collewijn, H. (1986). Necessary conditions for the perception of motion in depth. *Investigative Ophthalmology and Visual Science*, *27*, 584–597.
- Roberts, F. S. (1979). Measurement theory, with applications to decisionmaking, utility, and the social sciences. In G.-C. Rota (Ed.), *Encyclopedia of mathematics and its applications* (Vol. 7). Reading, MA: Addison-Wesley.
- Rogers, B. J., & Graham, M. E. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*, 125–134.
- Rogers, B. J., & Graham, M. E. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science*, *221*, 1409–1411.
- Shannon, C. E. (1949). The mathematical theory of communication. In C. E. Shannon & W. Weaver (Eds.), *The mathematical theory of communication*. Urbana, IL: University of Illinois Press (Original work published 1948).
- Steingrimsson, R. (2009). Evaluating a model for global psychophysical judgments of brightness: I. Behavioral properties of summations and productions. *Attention, Perception, & Psychophysics*, *71*, 1916–1930.
- Steinman, R. M., Levinson, J. Z., Collewijn, H., & van der Steen, J. (1985). Vision in the presence of known natural retinal image motion. *Journal of the Optical Society A*, *2*, 226–233.
- Stevens, M., & Merilaita, S. (2009a). Animal camouflage: Current issues and new perspectives. *Philosophical Transactions of the Royal Society B*, *364*, 423–427.
- Stevens, M., & Merilaita, S. (2009b). Defining disruptive coloration and distinguishing its functions. *Philosophical Transactions of the Royal Society B*, *364*, 481–488.
- Stevens, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1–49). New York: Wiley.
- Stewart, I. (2007). *Why beauty is truth, a history of symmetry*. New York: Basic Books.

- Suppes, P., & Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 1–76). New York: Wiley.
- Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics*. Mahwah, NJ: Erlbaum.
- Tadin, D., Haglund, R. F., Jr., Lappin, J. S., & Peters, R. A. (2001). Effects of surface microstructure on macroscopic image shading. *Proceedings of SPIE Conference on Human Vision and Electronic Imaging VI*, 4299, 221–230.
- Tankus, A., & Yeshurun, Y. (2009). Computer vision, camouflage breaking and countershading. *Philosophical Transactions of the Royal Society B*, 364, 529–536.
- Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3D objects in luminance noise. *Vision Research*, 35, 3053–3069.
- Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Sciences*, 8, 115–121.
- Todd, J. T., Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1996). Effects of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 695–706.
- Todd, J. T., & Norman, J. F. (2003). The visual perception of 3-D shape from multiple cues: Are observers capable of perceiving metric structure? *Perception & Psychophysics*, 65, 31–47.
- Todd, J. T., Norman, J. F., Koenderink, J. J., & Kappers, A. M. (1997). Effects of texture, illumination, and surface reflectance on stereoscopic shape perception. *Perception*, 26, 807–822.
- Todd, J. T., & Oomes, A. H. J. (2002). Generic and non-generic conditions for the perception of surface shape from texture. *Vision Research*, 42, 837–850.
- Todd, J. T., Oomes, A. H. J., Koenderink, J. J., & Kappers, A. M. L. (2004). The perception of doubly curved surfaces from anisotropic textures. *Psychological Science*, 15, 40–46.
- Todd, J. T., & Reichel, F. D. (1990). The visual perception of smoothly curved surfaces from double-projected contour patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 665–674.
- Todd, J. T., Thaler, L., & Dijkstra, T. M. H. (2005). The effects of field of view on the perception of 3D slant from texture. *Vision Research*, 45, 1501–1517.
- Todd, J. T., Thaler, L., Dijkstra, T. M. H., Koenderink, J. J., & Kappers, A. M. L. (2007). The effects of viewing angle, camera angle, and sign of surface curvature on the perception of three-dimensional shape from texture. *Journal of Vision*, 9, 1–16. doi:10.1167/7.12.9. 7(12).
- Troscianko, T., Benton, C. P., Lovell, P. G., Tolhurst, D. J., & Pizlo, Z. (2009). Camouflage and visual perception. *Philosophical Transactions of the Royal Society B*, 364, 449–461.
- Turner, J., Braunstein, M. L., & Andersen, G. J. (1995). Detection of surfaces in structure from motion. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 809–821. doi:10.1037/0096-1523.21.4.809
- Tyler, C. W. (1971). Stereoscopic depth movement: Two eyes less sensitive than one. *Science*, 174, 958–961.
- Tyler, C. W. (1973). Periodic vernier acuity. *Journal of Physiology (London)*, 228, 637–647.
- Uttal, W. R. (1975). *An autocorrelation model of form detection*. Hillsdale, NJ: Erlbaum.
- van Doorn, A. J., Koenderink, J. J., & Wagemans, J. (2011). Light fields and shape from shading. *Journal of Vision*, 21, 1–12. doi:10.1167/11.3.21. 11(3).
- van Ee, R., Adams, W. J., & Mamassian, P. (2003). Bayesian modeling of cue interaction: Bistability in stereoscopic slant perception. *Journal of the Optical Society of America. A*, 20, 1398–1406.
- van Ee, R., & Erkelens, C. J. (1996). Stability of binocular depth perception with moving head and eyes. *Vision Research*, 36, 3827–3842.
- Vuong, Q. C., Domini, F., & Caudek, C. (2006). Disparity and shading cues cooperate for surface interpolation. *Perception*, 35, 145–155.
- Wagemans, J., De Winter, H., Op de Beeck, H., Ploeger, A., Beckers, T., & Vanroose, P. (2008). Identification of everyday objects on the basis of silhouette and outline versions. *Perception*, 37, 207–244.
- Wagemans, J., van Doorn, A. J., & Koenderink, J. J. (2010). *The shading cue in context*. *iPerception*, 1, 159–178.
- Wagemans, J., van Doorn, A. J., & Koenderink, J. J. (2011). Measuring 3D point configurations in pictorial space. *iPerception*, 2, 77–111.
- Watt, R. J., & Andrews, D. P. (1982). Contour curvature analysis: Hyperacuties in the discrimination of detailed shape. *Vision Research*, 22, 449–460.
- Westheimer, G. (1975). Visual acuity and hyperacuity. *Investigative Ophthalmology*, 14, 570–572.
- Westheimer, G. (1977). Spatial frequency and light-spread descriptions of visual acuity and hyperacuity. *Journal of the Optical Society of America*, 67, 207–212. doi:10.1364/JOSA.67.000207
- Westheimer, G. (1979). The spatial sense of the eye: Proctor lecture. *Investigative Ophthalmology*, 18, 893–912.
- Westheimer, G., & McKee, S. P. (1978). Stereoscopic acuity for moving retinal images. *Journal of the Optical Society of America*, 68, 450–455. doi:10.1364/JOSA.68.000450
- Wiener, N. (1954). *The human use of human beings: Cybernetics and society*. Boston: Houghton Mifflin.
- Wiener, N. (1961). *Cybernetics, or control and communication in the animal and the machine* (2nd ed.). Cambridge, MA: MIT Press.
- Wilkinson, F., Wilson, H. R., & Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, 38, 3555–3568. doi:10.1016/S0042-6989(98)00039-X
- Wilson, H. R. (1985). Discrimination of contour curvature: data and theory. *Journal of the Optical Society of America. A*, 2, 1191–1198.
- Wilson, H. R., & Wilkinson, F. (1998). Detection of global structure in Glass patterns: Implications for form vision. *Vision Research*, 38, 2933–2947. doi:10.1016/S0042-6989(98)00109-6
- Zaidi, Q., & Li, A. (2002). Limitations on shape information provided by texture cues. *Vision Research*, 42, 815–835.