



Observer biases in the 3D interpretation of line drawings

Pascal Mamassian *, Michael S. Landy

Department of Psychology and Center for Neural Science, New York University, 6 Washington Place, 8th Floor, New York, NY 10003, USA

Received 6 May 1997; received in revised form 20 November 1997

Abstract

Line drawings produced by contours traced on a surface can produce a vivid impression of the surface shape. The stability of this perception is notable considering that the information provided by the surface contours is quite ambiguous. We have studied the stability of line drawing perception from psychophysical and computational standpoints. For a given family of simple line drawings, human observers could perceive the drawings as depicting either an elliptic (egg-shaped) or hyperbolic (saddle-shaped) smooth surface patch. Rotation of the image along the line of sight and change in aspect ratio of the line drawing could bias the observer toward either interpretation. The results were modeled by a simple Bayesian observer that computes the probability to choose either interpretation given the information in the image and prior preferences. The model's decision rule is noncommitting: for a given input image its responses are still probabilistic, reflecting variability in the modeled observers' judgements. A good fit to the data was obtained when three observer assumptions were introduced: a preference for convex surfaces, a preference for surface contours aligned with the principal lines of curvature, and a preference for a surface orientation consistent with an object viewed from above. We discuss how these assumptions might reflect regularities of the visual world. © 1998 Elsevier Science Ltd. All rights reserved.

Keywords: Line drawings; Surface contours; Geodesic curves

1. Introduction

Several cues in the image enable us to estimate the 3D structure of our environment [1]. For instance, the depth, orientation, and shape of surfaces can be computed from the disparities between the two eyes, the relative motion of object features or the patterns of shading. Each cue focuses on only one aspect of the image and processes this information to compute one attribute of the environment. Due to the limited amount of information available from an image, these computations are often under-constrained, i.e. there is not always a unique solution to the computational problem. For instance, to use the binocular disparities, one has to overcome the problem of multiple correspondences between the features in the left and right images [2]. Similarly, the inference of shape from shading is unwarranted unless one also knows the illumination conditions [3].

There are two fundamental ways in which the visual system can overcome the under-constrained problem. One way is to take advantage of several cues to fill in the missing parameters of individual cues. For instance, binocular horizontal disparities cannot be converted to absolute depth values without an estimate of the viewing distance, but this distance can be obtained, for example, from motion parallax. The process by which a cue interacts with others to provide absolute depth is known as cue promotion [4].

The other way one can constrain the computations of a depth cue is to impose a priori assumptions on the missing parameters. For instance, assuming that the surface is smooth helps one solve the correspondence problem for stereopsis [2]. One example where prior knowledge plays a critical role is for the interpretation of simple pictures. When an observer views a picture or photograph, the pictorial cues to depth indicate a 3D scene totally at variance with other cues such as stereopsis or motion parallax (the latter cues signal the true situation: the scene is flat). Human observers have a dual perception in such a situation and can make judgements of the depicted 3D scene as well as of the

* Corresponding author. Present address: Department of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QQ, Scotland, UK. E-mail: pascal@psy.gla.ac.uk.

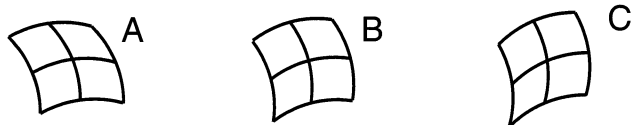


Fig. 1. The three figures used in the experiment varied in aspect ratio. At most orientations in the image plane, (A) was more frequently perceived as egg-shaped and (C) as saddle-shaped. (B) was the most ambiguous.

actual flat surface [5–7]. Consider for instance the line drawings in Fig. 1. Even though these images are very impoverished depth stimuli, consistent with an infinite number of 3D scenes, we shall see that human observers are considerably robust in their interpretation. That only a small number of distinct percepts result from each drawing indicates that the observer applies some assumptions (or prior constraints) while the contours are processed. The present investigation looks at these assumptions via the biases that observers exhibit in the interpretation of line drawings.

1.1. Shape-from-contour

Lines in an image can have their origin in a number of contours in the scene. One important family of contours, called surface contours, corresponds to changes in surface reflectance (surface markings) or from changes in illumination such as cast shadows [8–10]. As such, surface contours are less constrained than occluding contours which are formed at the boundary between the object and the background [11,12]. Surface contours also differ in principle from texture markings where each texture element is assumed to be small relative to the surface curvature (i.e. to lie approximately in a local tangent plane). Thus, the estimation of the shape of the surface from the surface contours (shape-from-contour) is essentially different from the estimation of shape from the texture markings (shape-from-texture) [13].

As an infinite number of 3D contours can give rise to a given image contour, the problem of shape-from-contour is severely underconstrained. In spite of this ambiguity, image contours lead to a 3D impression comparable to that obtained from other depth cues such as binocular disparities [14]. In an attempt to bridge the gap between the underconstrained nature of image contours and human ability to interpret them, Stevens proposed that an observer made three basic assumptions [15]. These assumptions were that: (1) surface contours were lines of curvature; (2) surface contours were geodesics; and (3) object and observer were in general position. Lines of curvature are the directions on a surface along which the local orientation changes the most and the least (on a cylinder, these curves are the parallels and the meridians). The use of

this assumption is consistent with the human interpretation of images where a surface is depicted by parallel contours [15,16]. Geodesic curves on a surface are generalizations of straight lines on a plane: their curvature is solely due to the surface curvature. Recent psychophysical studies have also shown the relevance of the assumption of geodesicity for human perception [17]. Finally, the general viewpoint assumption is critical to limit possible interpretations of a scene so that, for instance, a straight image contour is never interpreted as an accidental projection of a curved 3D contour. The general viewpoint assumption has recently received a renewal of interest [18,19].

1.2. Overview

In the present study, we shall take both a psychophysical and computational approach to uncover the set of assumptions that observers use to interpret line drawings. For this purpose, we created simple images composed of two sets of parallel contours which are typically perceived as 4-way ambiguous surface patches (Fig. 1). Each set of contours can appear either convex or concave, resulting in convex and concave egg-shaped interpretations and two saddle-shaped interpretations. Each of these interpretations occurs more or less frequently depending on the geometry of the figure. We propose that these biases of the visual system result from specific assumptions made by observers concerning the objects in the scene as well as the viewpoint. These assumptions are assessed with the help of a stochastic model that knows about the geometry of image formation and incorporates explicit prior constraints. By construction, this model is a particularly simple instance of the class of Bayesian models [20–22].

2. Human observers' biases

We investigated the perception of solid shape from line drawings. Local solid shape can be considered as a continuum from which two fundamental categories emerge: the elliptic shape (i.e. locally egg-shaped) and the hyperbolic shape (saddle-shaped) [23]. Observers viewed line drawings such as those shown in Fig. 1 and were required to categorize their first impression as elliptic or hyperbolic. Their biases for either interpretation were recorded for different image orientations and different aspect ratios of the figure.

2.1. Methods

2.1.1. Subjects

Seven human adults participated in this experiment. All had normal or corrected-to-normal visual acuity.

2.1.2. Stimuli

The stimuli were made out of six identical circular arcs (45° arcs) arranged in two sets of parallel contours (Fig. 1). The two ends of each arc were 3.8° apart, as viewed from a distance of 50 cm. Three figures were used in the experiment, their difference lying in their aspect ratio. Shape A was compressed along its axis of symmetry, while shape C was stretched along the axis of symmetry; shape B was an intermediate figure between A and C. These elongations were such that the central intersection of A, B and C formed an angle of 75 , 90 and 105° , respectively. The stimuli were presented on a 17" Sony Trinitron monitor controlled by a Macintosh 7100/80 computer running the PsyScope software [24]. Image resolution was 1024×768 pixels at a refresh rate of 75 Hz. Stimuli were displayed as black lines on a 77 cd/m^2 white background in a darkened room.

2.1.3. Procedure

The design followed a two alternative, one interval, forced-choice paradigm. On each trial, observers viewed a single figure and reported with the help of two keys on a keyboard whether their first impression was that of a saddle-shaped or egg-shaped surface. Before being displayed on the monitor, the figure was randomly oriented in the image plane in steps of 15° . A block of 72 trials consisted of each of the three figures at each of the 24 orientations in random order. Subjects completed 12 such blocks. Viewing was monocular, and viewing time was unrestricted. After each response, there was a blank interval and the next trial was started with a keypress.

2.2. Results and discussion

The results, pooled over the seven subjects, are shown in Fig. 2. In these polar plots, the angle represents the orientation of the figure in the image plane, and the radius the proportion of times the figure was perceived as being egg-shaped (the 'elliptic score'). The error bars represent the standard deviation of the means obtained across observers. Due to these standard errors usually being small, we shall only consider the pooled data rather than individual subject results.

The first thing to note is that the aspect ratio of the figure had a dramatic effect on the perceived shape. Shape A, which was compressed along its axis of symmetry, was perceived more often as egg-shaped (high elliptic score) in all orientations, whereas shape C was perceived more often as saddle-shaped (low elliptic score). The intermediate shape B was perceived differently depending on its orientation in the frontal plane: when its axis of symmetry was vertical, the figure was perceived more often as egg-shaped, whereas when this axis was close to the horizontal, the figure was per-

ceived as saddle-shaped. In other words, the interpretation of shape B could change from egg-shaped to saddle-shaped by a rotation of 90° in the frontal plane (Fig. 3).

The effect of image orientation on the 3D interpretation of line drawings was also briefly noted by Stevens [16]. However, there was nothing in Stevens' model [15] that could account for this image orientation effect, because only the curvature of the contours was relevant (contours were assumed to be principal lines of curvatures). One potential explanation for the orientation-dependence is that observers interpret those contours which are convex-upward in the image as originating from convex markings on the surface of the object (Fig. 4). When two contours intersect in the image, the line drawing is interpreted as being elliptic if both contours are convex or concave, and hyperbolic if one contour is convex and the other is concave. The effect of image orientation is subsequently predicted on the basis that convex contours become concave (and vice versa) when the image is rotated by 180° . The correspondence between convexity in the image and convexity of the surface is in fact an assumption that the surface is oriented in such a way that its normal points upward, as if the observer looked down at a ground plane on which the object lay (Fig. 5). This intuition will be made more precise with the help of the model developed in the next section.

Finally, it is important to note that the change of interpretation from egg-shaped to saddle-shaped for Shape B is gradual: there is no sharp categorization of the line drawings into egg-shaped and saddle-shaped surface patches. A model intended to account fully for the observers' performance should reflect this property of the data. Our approach to this problem is the definition of a 'simple Bayesian observer'.

3. Simple Bayesian observer

Given a particular line drawing stimulus, we wish to predict the probability that observers will perceive the surface as elliptic. Since the data are clearly driven by observer biases, a natural approach is to adopt the Bayesian framework. After giving an outline of the method, we detail every step that leads to the construction of a Bayesian model for our problem.

3.1. Model construction

3.1.1. Rationale

The fact that human observers were biased toward an elliptic or hyperbolic interpretation for a given line drawing suggests that they used some assumptions to interpret the image. Bayesian analysis is a natural approach for including these assumptions into the observ-

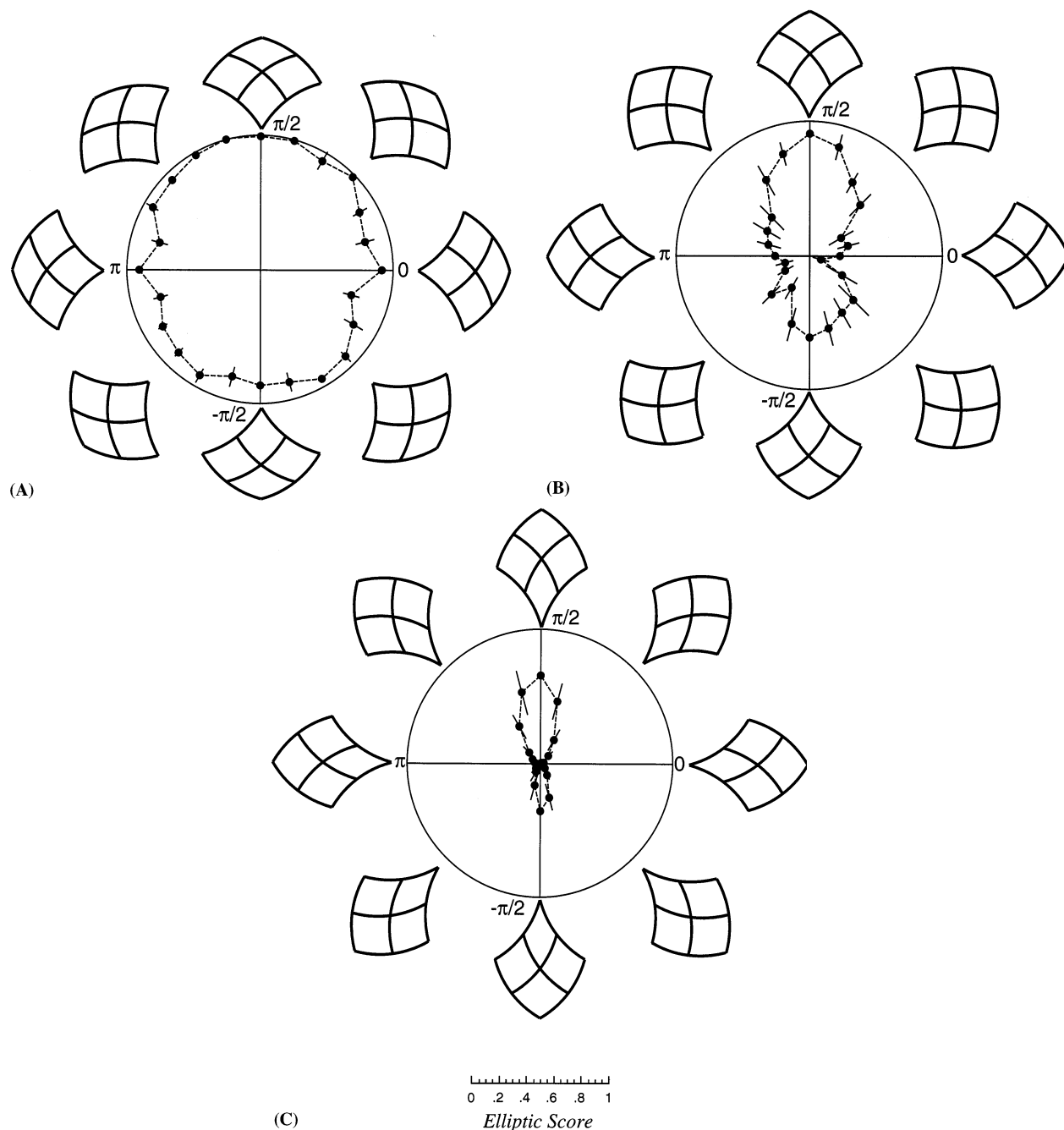


Fig. 2. Each panel shows the results of the experiment in a polar plot, averaged over the seven subjects, for the corresponding contour shape pictured in Fig. 1. Distance from the origin gives the percentage of occasions in which this shape in this orientation was perceived to be elliptical (egg-shaped). Shape A was nearly always perceived as egg-shaped; shape C was more often perceived as saddle-shaped, but the choice probability depended upon the orientation. Perception of shape B fell in between. The error bars represent plus-or-minus one standard error of the inter-subject variability.

er's inference process [25]. A model within this Bayesian framework is composed of two basic stages [26]. First, the assumptions are coded as prior probability distributions which are then combined with the information in the image using Bayes' rule. For our problem, this computation will lead to computation of the probability that the surface patch is elliptical for a given line drawing image (the posterior probability). The purpose

of the second stage of the model is to convert the posterior probability into an actual response of the observer ('it is elliptical' or 'it is hyperbolic') by following a specific decision rule.

There are a variety of ways to encode image measurements in the model. To keep the model tractable, we do not consider the entire line drawing, but merely the central intersection where the contours overlap to form

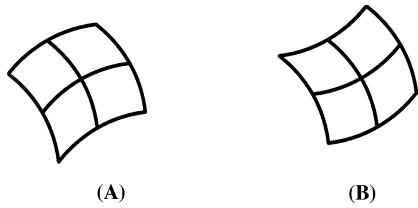


Fig. 3. The interpretation of shape B was dependent on its orientation in the frontal plane. In (A), it was perceived more often as egg-shaped, whereas in (B), it was perceived more often as saddle-shaped.

an X. However, due to the symmetry of the line drawing, we expect that the contribution of another intersection will be counterbalanced by the intersection symmetrically opposite, and thus we do not think that this simplification is critical.

Due to the curvature of the contours being held constant in our psychophysical experiment, we shall not consider the degree of curvature in the present model. Only the sign of curvature is taken into account to discriminate convex from concave contours. For each line drawing stimulus, we therefore measure the orientation of each of the two curves where they intersect, and code this orientation relative to the normal to the curves. Thus, these two orientations φ_1 and φ_2 lie between 0 and π if the curve is locally convex-upward,

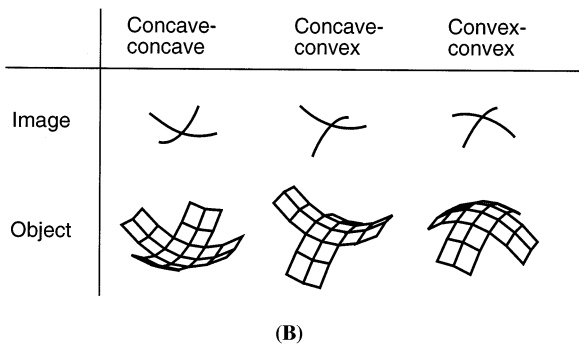
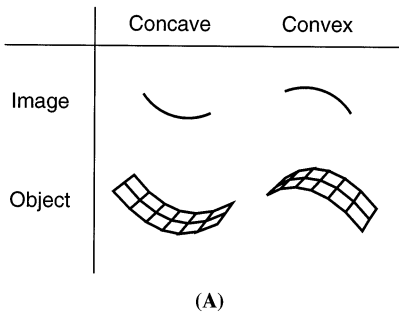


Fig. 4. (A) Convex contours in the image can be assumed to be produced by convex contours on a surface (and similarly for concave contours). (B) In this case, elliptic surface patches will be perceived when the contours are both convex or both concave in the image, and hyperbolic patches will be seen when one contour is convex and the other is concave.

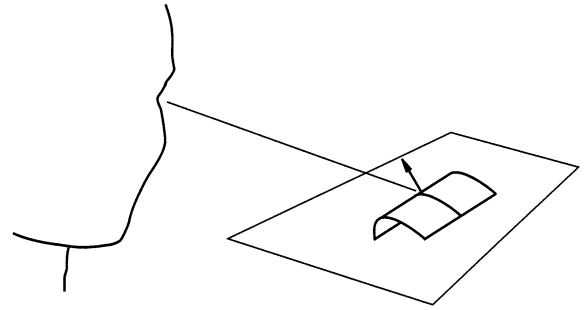


Fig. 5. The bias on the tilt of the surface is consistent with a bias on the relative positions of the observer and the object, as if the observer were above the object (e.g., the object lay on the ground plane).

and between π and 2π , if it is convex-downward (Fig. 6). Using this notation, the posterior probability to be computed by the first stage of the model can be written

$$P(\text{elliptic} | \varphi_1, \varphi_2), \tag{1}$$

in other words, the posterior probability is the probability that the surface is elliptic given that the orientation of the two curves intersecting at the center of the line drawing is measured to be φ_1 and φ_2 .

To compute this probability, we shall have to estimate the likelihood that a certain scene (a painted 3D patch at a certain position) gives rise to the image data (described by φ_1 and φ_2), and model the prior probability for each scene (before having looked at the image). Geometrically speaking, the scene is characterized by three basic factors: (1) the orientation of the patch relative to the observer; (2) the local shape at the center of the patch; and (3) the orientation of the surface markings on the patch. We shall provide a parameterized stochastic model (in the form of a prior distribution) for each of these three factors. More specifically, we shall characterize a preference for perceiving a surface (1) oriented such that its normal points upward, (2) locally convex, and (3) such that the surface contours are aligned with the principal lines of curvature. We

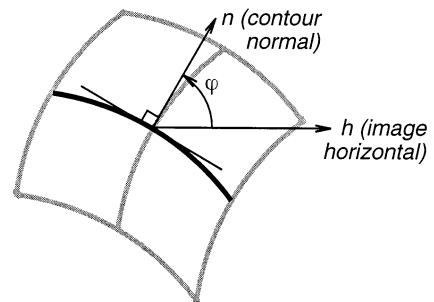


Fig. 6. Definition of φ for the two contours intersecting at the center of the figure. The image contour orientation φ is the angle in the image plane that the normal to the contour makes with the horizontal axis. As such, it takes into account both the orientation of the contour tangent as well as the sign of the image plane contour curvature.

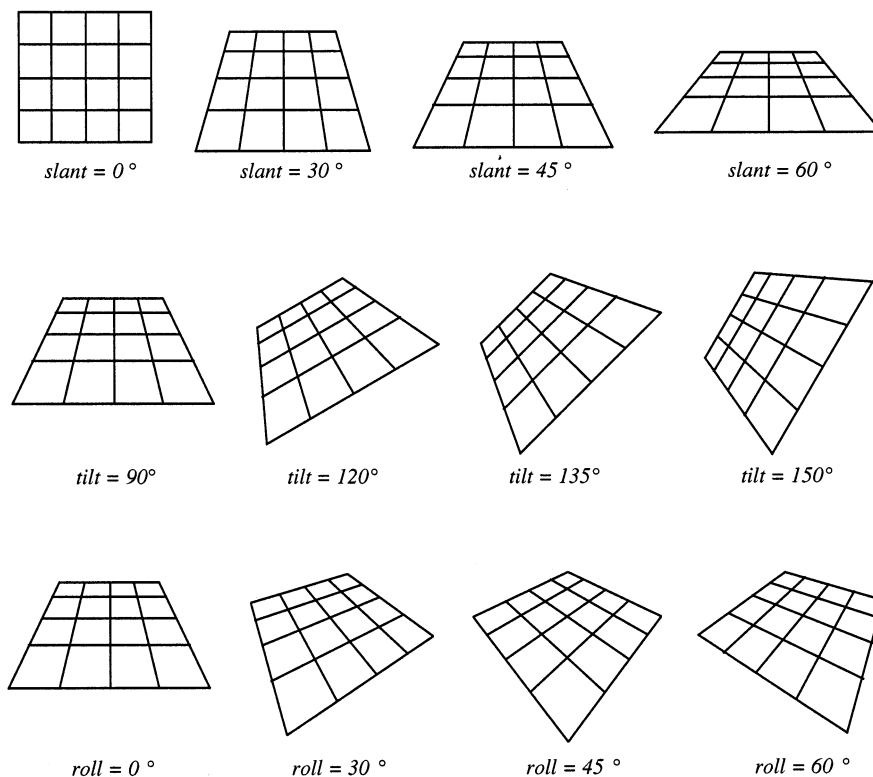


Fig. 7. Orientation of a surface patch in three-dimensional space. The slant accounts for the inclination of the object relative to the line of sight (first row). The tilt indicates the direction of steepest descent (second row). We choose a coordinate frame aligned with the principal curvatures. The roll refers to the rotation about the surface normal that relates the direction of steepest descent to the direction of the first principal curvature. Within each row, only one angle varies; the two others are set to the following default values: slant = 45° , tilt = 90° , roll = 0° .

shall see that these three biases suffice to explain our psychophysical results, so that for all the other dimensions of the model, uniform prior distributions will be taken by default. Then, we shall assume that surface orientation, shape and surface marking orientation constitute independent events so that the posterior probability can be simply derived using Bayes' rule. Finally, we shall have to choose a decision rule to map the posterior probability onto an elliptic score for the model (i.e. the proportion of times the model would choose 'elliptic' rather than 'hyperbolic' for a given line drawing). We shall argue that the particular conditions of our psychophysical experiment are compatible with the simplest of all decision rules so that we shall call our model a 'simple Bayesian observer'.

In the following sections, we shall derive, one by one, the three prior distributions for the three geometrical factors outlined above, and compute the likelihood function for our problem. We shall then combine the prior distributions with the likelihood function using Bayes' rule, apply our decision rule and adjust the parameters in the model so as to fit the empirical data.

3.1.2. Surface orientation prior distribution

The orientation of a surface patch can be conveniently represented by three angles, the tilt τ , the slant

σ and the roll ρ (Fig. 7). The tilt indicates the image plane orientation of the surface normal (or equivalently of the surface gradient), and thus ranges from 0 to 2π . The slant measures the degree to which the surface is rotated away from fronto-parallel, and hence ranges from 0 to $\pi/2$. The roll is a self rotation of the surface patch in its tangent plane, and thus also ranges from 0 to 2π . We now propose prior distributions for τ , σ and ρ .

If the observer does not have any bias on perceived tilt, the image should always be interpreted the same way independent of its orientation in the frontal plane. In other studies [27], we have shown that this is not the case: the interpretation of a line drawing is affected by

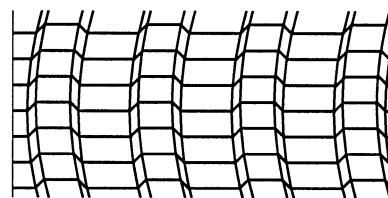


Fig. 8. An embossed surface defined by shape-from-contour. The thin stripes seen in relief here are perceived as indentations if the image is rotated by 180° in the image plane, illustrating the visual system's assumption that the viewpoint is located above the object.

a rotation of the image in the frontal plane. For instance, rotating Fig. 8 by 180° changes the interpretation from a surface with narrow ridges to a surface with narrow troughs. We conclude from this phenomenon that observers had a bias on surface orientation, namely a preference for surfaces whose surface normals point upward rather than downward. When the surface is roughly horizontal (such as when an object lies on a ground plane), this bias takes a very intuitive form: surfaces are more often seen from above rather than from below (Fig. 5).

This bias on surface tilt is also consistent with the results of the experiment described here. Thus, we shall allow for a bias of the surface tilt in the $\pi/2$ or upward direction. For simplicity, we assume a Gaussian distribution (due to orientation being periodic, we should use, in theory, the formalism of circular statistics [28]; in practice, however, a Gaussian distribution will be acceptable as long as its variance is small):

$$P(\tau) = \frac{1}{\sqrt{2\pi\sigma_\tau}} \exp\left(-\frac{(\tau - \pi/2)^2}{2\sigma_\tau^2}\right). \quad (2)$$

The standard deviation of the Gaussian, σ_τ , controls the degree of bias. A value of $\sigma_\tau = \infty$ corresponds to no bias for the viewing direction.

We have no evidence for any observer biases concerning surface slant, therefore, we shall assume an unbiased prior distribution. More precisely, we assume a distribution of slant σ that would lead to a uniform distribution of surface orientation if tilt were also unbiased (i.e. equal areas on the Gaussian sphere are equally likely). A simple calculation results in the prior distribution

$$P(\sigma) = \sin(\sigma). \quad (3)$$

Likewise, we have no evidence for any observer biases concerning surface roll, therefore, we shall assume an unbiased prior distribution. We define surface roll relative to the direction of the steepest descent of the surface. Intuitively, changing surface roll corresponds to rotating the surface relative to its normal. To measure the amount of such a rotation, we need a distinguished direction on the surface to serve as a reference. This direction is naturally provided by one of the principal directions—the directions along which the normal sections to the surface have largest and smallest curvatures [29,30]. We define surface roll ρ as the angle between the direction of steepest descent and the first principal direction (the second principal direction will be oriented at $\rho + \pi/2$ by definition since the principal directions are always orthogonal). Due to the local surface shape being π -periodic, and due to the principal directions playing interchangeable roles, we can restrict the distribution of ρ to be uniform between 0 and $\pi/2$.

Assuming that the three components of surface orientation are mutually independent, we can deduce the

joint probability $P(\tau, \sigma, \rho)$ as the product of the individual probabilities.

3.1.3. Local shape prior distribution

Solid shape can be locally summarized by two parameters such as the two principal curvatures of the surface, k_1 and k_2 . Mamassian et al. [23] favored a different parameterization of shape, based on two parameters they termed the shape characteristic and curvature magnitude. Shape characteristic χ indicates whether the surface is locally elliptic ($0 < \chi \leq 1$) or hyperbolic ($-1 \leq \chi < 0$), while the curvature magnitude η indicates whether the surface is locally convex ($\eta > 0$) or concave ($\eta < 0$). These two entities are related to the principal curvatures as follows:

$$\begin{cases} \chi = k_1/k_2 \\ \eta = -k_2 \end{cases} \quad \text{with } |k_1| \leq |k_2|. \quad (4)$$

Subjects were asked to judge whether the perceived shape was elliptic or hyperbolic. Whether the surface is convex or concave is not relevant in this task. However, there are some indications that convex and concave surfaces are not equivalently treated by the observers. This can be appreciated by looking again at Fig. 2(A): the line drawing placed at the top of the plot is usually perceived as being convex, whereas the same line drawing rotated by 180° (at the bottom of the plot) is more often perceived as being concave. We propose that the difference in interpretation between concave and convex is related to the difference in proportion of elliptic responses for these two figures.

How can we account for such a difference between convex and concave surfaces in our model? The convexity of the surface is represented by the curvature magnitude, where the latter is simply proportional to one of the principal curvatures. Therefore, a natural way to allow for a bias on the surface convexity is to posit a bias on the individual principal curvatures. For no particular reason, we choose this bias to follow a Gaussian distribution:

$$P(k_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(k_i - \mu_k)^2}{2}\right), \quad i = 1, 2. \quad (5)$$

In our characterization of the line drawing, we have paid attention only to the sign of the image curvature at the X-junction, not to the amount of curvature itself. This is an important simplification that was made possible by not varying the image plane curvature of the line segments. As we consider only the difference between convex and concave segments, one might think that we could have simply used a binomial distribution for the principal curvature rather than the above continuous distribution. However, we are going to see (Eq. (12)) that a continuous distribution is necessary for the computation of the likelihood function. The only relevant feature of the above bias is the probability that

each principal curvature is positive (or negative), not the details of the distribution. Thus, we have arbitrarily set the standard deviation of the distribution to 1, and the single parameter μ_k suffices. A negative value of μ_k corresponds to a bias in favor of contours on a convex surface.

Using Eq. (4), one can deduce the prior probability on surface shape $P(\chi, \eta)$ from the biases for k_1 and k_2 .

3.1.4. Surface marking orientation prior distribution

A remarkable feature of the data in the current study is how a minor change in the aspect ratio of the figure produces such a large bias in the perceived shape (compare again shapes A and C). In developing the model, we found this was a very difficult feature of the data to reproduce. There is however one aspect of the scene that we have not yet exploited, namely the way the surface markings are arranged on the 3D patch. We first tried to favor configurations where the two contours were geodesic and orthogonal to each other (that is two orthogonal contours whose curvature is entirely the result of the surface curvature). However, as we will see in the forthcoming section dealing with simulations of the model, this bias was not sufficient to account for the big shift in perceived shape. A stronger constraint is that proposed by Stevens [15] where he assumed that surface markings are oriented approximately along the principal directions (which form a particular pair of orthogonal geodesic curves). Again for simplicity, we choose this bias to follow a Gaussian distribution centered on each of the principal directions. Denoting by ψ the orientation of a surface marking in the tangent plane relative to the first principal direction, we have

$$\begin{cases} P(\psi_1) = \frac{1}{\sqrt{2\pi\sigma_\psi}} \exp\left(-\frac{(\psi_1 - \rho)^2}{2\sigma_\psi^2}\right) \\ P(\psi_2) = \frac{1}{\sqrt{2\pi\sigma_\psi}} \exp\left(-\frac{(\psi_2 - \rho - \pi/2)^2}{2\sigma_\psi^2}\right) \end{cases} \quad (6)$$

The standard deviation σ_ψ controls the strength of this bias. We assume that the deviation of the first surface marking from the first principal direction is independent of the deviation of the second surface marking from the second principal direction. However, since the two principal directions are always orthogonal to each other, the orientations of the two surface markings are not strictly independent. It is nevertheless possible to compute the joint probability $P(\psi_1, \psi_2)$.

3.1.5. Bayesian combination

We have characterized the scene in terms of surface orientation (τ, σ and ρ), local surface shape (η and χ) and surface markings (ψ_1 and ψ_2) and provided parameterized probability distributions for each of these variables. Our model computes the posterior

probabilities in terms of these variables, given the image orientations φ_1 and φ_2 . Using the principle of total probability, the posterior probability can be expanded as

$$\begin{aligned} & P(\text{elliptic}|\varphi_1, \varphi_2) \\ &= \int_{\Omega} P(\tau, \sigma, \rho, \chi, \eta, \psi_1, \psi_2|\varphi_1, \varphi_2) \\ & \quad \times d\tau d\sigma d\rho d\chi d\eta d\psi_1 d\psi_2. \end{aligned} \quad (7)$$

The integration takes place over the domain Ω which is the product of the domains for surface orientation, surface shape, and surface marking orientations. The domain for surface orientation spans τ over $(-\pi, \pi)$, σ over $(0, \pi/2)$ and ρ over $(-\pi/2, \pi/2)$. The shape is elliptic when χ is positive, therefore, the domain for surface shape spans χ over $(0, 1)$ and η over $(-\infty, \infty)$. Finally, the variables ψ_1 and ψ_2 account for the surface marking orientations which vary over $(-\pi/2, \pi/2)$.

Using Bayes' rule, we can rewrite the integrand of Eq. (7) as

$$\frac{P(\varphi_1, \varphi_2|\tau, \sigma, \rho, \chi, \eta, \psi_1, \psi_2)P(\tau, \sigma, \rho, \chi, \eta, \psi_1, \psi_2)}{P(\varphi_1, \varphi_2)}. \quad (8)$$

We have described parameterized observer priors for the surface orientation $P(\tau, \sigma, \rho)$, for the local surface shape $P(\chi, \eta)$, and for the surface marking orientations $P(\psi_1, \psi_2)$. Assuming that these three factors are independent, the integrand in Eq. (7) becomes

$$\frac{P(\varphi_1, \varphi_2|\tau, \sigma, \rho, \chi, \eta, \psi_1, \psi_2)P(\tau, \sigma, \rho)P(\chi, \eta)P(\psi_1, \psi_2)}{P(\varphi_1, \varphi_2)}. \quad (9)$$

3.1.6. Likelihood

In the expression just obtained (Eq. (9)), the first term in the numerator is called the likelihood. This term represents the image formation process: is the image (again restricted to the X-junction) compatible with the projection of a particular 3D scene (i.e. a particular orientation of a surface patch with two painted lines)? If the image formation is noise-free (that is, if the projection of an object to the image plane is uniquely determined) then this likelihood term is a delta function

$$\begin{cases} P(\varphi_1, \varphi_2|\tau, \sigma, \rho, \chi, \eta, \psi_1, \psi_2) = \\ \left\{ \begin{array}{l} 1, \quad \text{if } \varphi_1 = F(\tau, \sigma, \rho, \chi, \eta, \psi_1) \text{ and } \varphi_2 = F(\tau, \sigma, \rho, \chi, \eta, \psi_2) \\ 0 \quad \text{otherwise} \end{array} \right. \end{cases} \quad (10)$$

where $F(\tau, \sigma, \rho, \chi, \eta, \psi)$ is the function which projects a curve at orientation ψ on the surface onto a curve at orientation φ in the image (Fig. 9). Leaving the geometrical derivation of this function to Appendix A, we find

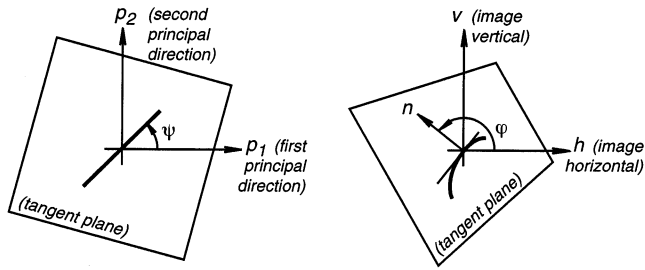


Fig. 9. Projection of the curve orientation in the tangent plane to a contour in the image. On the left, we are viewing the surface patch along its surface normal. Due to the curvature of the contour deriving entirely from the surface curvature, from this viewpoint the contour projects to a straight line (at least locally). The model begins by specifying the orientation ψ of the contour relative to the principal directions on the surface. The process of projecting this curve onto the image plane involves rotations corresponding to the roll, slant and tilt of the surface patch (Fig. 7). In the image plane, we describe the orientation ϕ of the normal to the contour relative to the horizontal axis.

$$F(\tau, \sigma, \rho, \chi, \eta, \psi) = \arctan\left(\frac{\tan(\psi + \rho)}{\cos(\sigma)}\right) + \text{sign}(k \tan(\psi + \rho)) \frac{\pi}{2} + \tau \quad (11)$$

where k is the curvature of the surface marking at the X-junction. If we assume that all of the curvature of the line is due to the curvature of the surface, i.e. the line is locally geodesic [17], then k can be written as a function of the two principal curvatures (the Euler formula; [31])

$$k = k_1 \cos^2(\psi) + k_2 \sin^2(\psi). \quad (12)$$

From Eq. (5), we know the probability density function of each principal curvature and can therefore compute the probability density of the curvature of each surface marking at the X-junction.

We are now able to compute our posterior probability in Eq. (9) up to a constant (namely, $P(\phi_1, \phi_2)$), which came from the Bayesian expansion. We actually do not need to compute this constant explicitly, since we can use the fact that

$$P(\text{elliptic}|\phi_1, \phi_2) + P(\text{hyperbolic}|\phi_1, \phi_2) = 1. \quad (13)$$

3.1.7. Decision rule

The computation of the posterior probability corresponds to the first stage of our Bayesian model. We now need to describe how observers use this probability to make an actual response when they view a given image. Several decision rules are commonly used in statistical decision theory [32,26]. One popular choice is the maximum a posteriori (MAP) decision rule whereby the observer chooses the interpretation with the highest posterior probability, as this decision rule minimizes the expected number of errors.

The choice of the appropriate decision rule for a problem involves a determination of the ‘cost’ of mak-

ing wrong decisions [32]. However, it is unjustifiable to apply this approach in the case of our experiment because the observer’s decision has no consequences and no feedback is given. Even though the inference of the local solid shape from the intersection of image contours can certainly play a role in the recognition of the line drawing of complex objects, this inference has no consequences in our laboratory conditions. In a sense, our observers might be facing a pure problem of statistical inference, that is to provide a summary of the statistical evidence for the estimated local shape with no further application. As such, it is not inappropriate for the observers’ judgements to reflect the computed posterior probability (‘probability matching’). In other words, the probability that the observer responds elliptic is the posterior probability for an elliptic shape (i.e. the probability we have just computed):

$$P(\text{Observer responds ‘elliptic’}|\text{image}) = P(\text{Surface shape is ‘elliptic’}|\text{image}). \quad (14)$$

We shall call a model for which the observer’s choice probability is the posterior probability a simple Bayesian model. There is an important difference between such a model and one linked with a more classical decision rule. Classical models often rely on a source of noise to account for the subject’s response variability (photon noise or internal noise in the case of visual psychophysics). In contrast, our model predicts the distribution of the observer’s responses solely from the posterior distribution. This posterior distribution reflects the ambiguity that multiple scenes can project to the same image, even when no source of noise is added.

3.2. Results

We were unable to find a closed-form solution for the posterior probability. Numerical simulations were run instead, where the Monte Carlo method was used to sample the prior distributions and build the posterior probability (Appendix B). The model has three parameters controlling the strength of the biases on surface orientation (parameter σ_τ), surface shape (μ_k), and surface marking orientation (σ_ψ), which we group into a vector Θ_3 . These parameters were adjusted so as to provide the maximum likelihood fit to the observers’ performance measured for three shapes and 24 orientations, i.e. 72 stimulus configurations (Fig. 2). The natural logarithm of this likelihood is (dropping terms that do not depend on the response data):

$$L(\Theta_3) = n \sum_{i=1}^{72} \{P_{\text{data}}(i) \text{Log}(P_{\text{model}}(i)) + (1 - P_{\text{data}}(i)) \text{Log}(1 - P_{\text{model}}(i))\}, \quad (15)$$

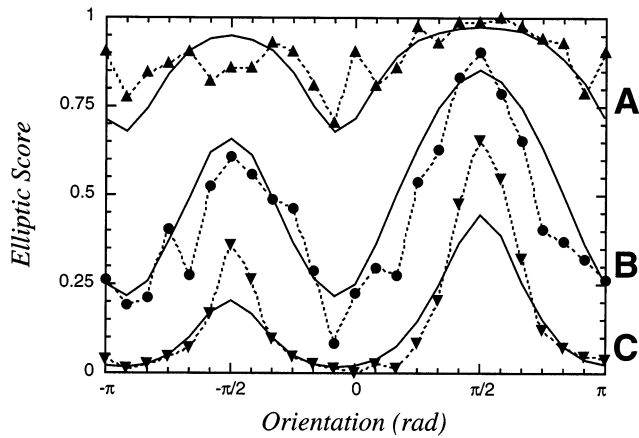


Fig. 10. The best fit obtained with the model is shown together with the data averaged over 7 subjects. The performance of the model is shown as solid lines. Although the fit is not perfect, it captures several important aspects of the data including the strong effects of shape and orientation, the asymmetry across 180° rotations, and the largely different interpretation of shapes A and C. Here, we switch from a polar to a Cartesian representation to avoid the bunching up of the data points visible in Fig. 2.

where $P_{\text{data}}(i)$ and $P_{\text{model}}(i)$ are the elliptic scores for stimulus i measured experimentally and predicted by the model, respectively. In this equation, n is the number of replications per stimulus configuration, that is 84 (12 replications for each of the 7 subjects). The best fit was obtained with parameter settings of $\sigma_\tau = 0.55$ rad; $\mu_k = -0.35$; and $\sigma_\psi = 0.13$ rad, and resulted in a logarithm of the likelihood $L(\Theta_3) = -2604$. The best fit is shown along with the empirical data in Fig. 10.

A value of -0.35 for μ_k corresponds to a ratio of convex to concave contours equal to 1.75 (i.e. a bias for convex segments) and a ratio of elliptic to hyperbolic surfaces equal to 1.16 (i.e. a slight bias for elliptic patches). The prior distributions associated with σ_τ and

σ_ψ are shown in Fig. 11. These are rather strong biases on the surface orientation and the position of the surface markings relative to the principal curvatures.

While deriving our model in the previous section, we wondered whether it would suffice to impose a constraint on the surface marking orientation to be simply geodesic and approximately orthogonal rather than approximately aligned with the principal directions. The modified model derived by imposing the orthogonal directions constraint can be simply obtained from the original model where we imposed the principal directions constraint. In Eq. (6), we just need to have ψ denote the orientation of a surface marking relative to an arbitrary direction rather than relative to the first principal direction. Simulations of the modified model produced a best fit that was notably worse than the original model (the logarithm of the likelihood was -2748 rather than -2604). The best fit of the modified model is shown along with the empirical data in Fig. 12.

In order to appreciate the role of each degree of freedom, we can look at a model deprived of a specific bias. For this purpose, one of the three parameters σ_τ , σ_ψ or μ_k was set to a value so as to eliminate the associated bias ($+\infty$, $+\infty$ and 0 , respectively). The remaining parameters of such deprived models were estimated again by maximum likelihood. These parameters for each model are summarized in Table 1, and the performance of the models is presented in Fig. 13. A bias on surface orientation (σ_τ) is required for image orientation to produce any effect on perceived shape because the other two biases are isotropic (they are intrinsic to the surface and surface markings, not to the particular image orientation). The bias on surface marking orientation (σ_ψ) is responsible for the massive difference in perception between the three shapes. Indeed, when surface markings are approximately aligned

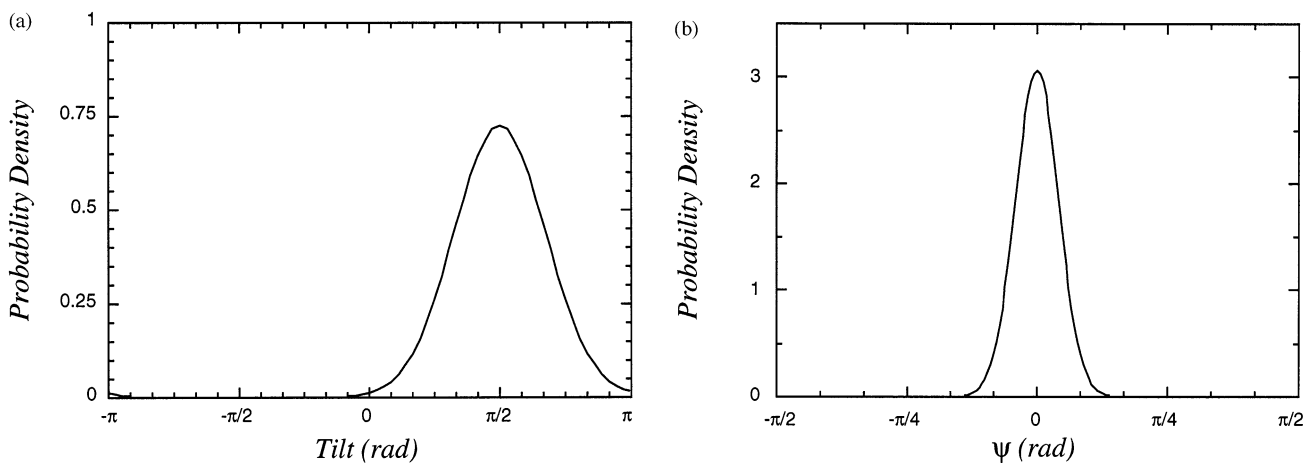


Fig. 11. The prior distributions corresponding to the best fitting model: (a) tilt distribution; (b) distribution of ψ_1 (the distribution of ψ_2 has the same shape, however, it is shifted by 90°). Both priors are sharply tuned. Observers are rarely willing to interpret surface contours as if the surface normal pointed downward, and are strongly biased to interpret the surface contours as closely aligned with the principal lines of curvature.

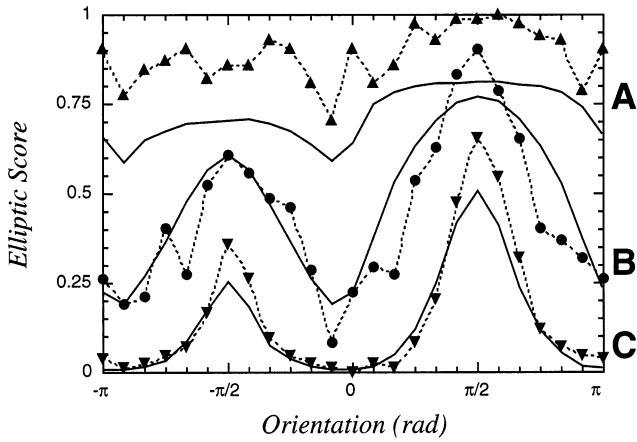


Fig. 12. The best fit obtained with the modified model where the constraint on surface markings was merely that they be geodesics and approximately orthogonal to each other in space. Comparing this fit with the one obtained with the original model (Fig. 10), we note that Shape A was less often judged ‘elliptic’.

with the principal directions, the sign of curvature of the two surface markings will determine the qualitative shape of the surface. In particular, this constraint prevents the possibility of a hyperbolic surface having two convex surface markings (that meet at an angle less than $\pi/2$ on the surface). Finally, the bias on surface shape (μ_k) is responsible for the asymmetries between the data at $\pi/2$ rad and $3\pi/2$ rad. Intuitively, this prior biases concave contours to be labeled as convex, resulting in more hyperbolic interpretations at $\pi/2$.

The contribution of each degree of freedom can be further assessed by testing the degree to which a deprived model fit the data less well than the original model. The significance of this loss in the fit can be evaluated using the nested hypothesis test [33] in which the null hypothesis is that the original model and a deprived one fit the experimental data equally well. Denoting by $L(\Theta_3)$ and $L(\Theta_2)$ the natural logarithm of the likelihood for the original (3 parameter) and deprived (2 parameter) models respectively (Eq. (15)), the test involves the statistic

$$2(L(\Theta_3) - L(\Theta_2)) \tag{16}$$

which is distributed as χ^2 with one degree of freedom (the difference in the number of model parameters).

Table 1
The best fit parameters for the Bayesian models developed to account for the subjects’ data

Model parameters	σ_τ	μ_k	σ_ψ	Log (likelihood)	χ^2 (1)	<i>p</i> -value
$(\sigma_\tau, \mu_k, \sigma_\psi)$	0.55	-0.35	0.13	-2604		
(μ_k, σ_ψ)	$+\infty$	-0.10	0.16	-2994	390	<0.001
$(\sigma_\tau, \sigma_\psi)$	0.58	0	0.14	-2689	85	<0.001
(σ_τ, μ_k)	0.56	-0.30	$+\infty$	-3286	682	<0.001

The models with two parameters are tested against the original model (with three parameters) for any significant loss in goodness of fit. All three two-parameter models are rejected.

These tests are summarized in Table 1. All models deprived of one parameter are handily rejected.

4. Discussion

We have investigated the human perception of simple line drawings. These images were interpreted as curved surface patches with contours painted on them. Even though there was an infinite number of curves in 3D that could project to these line drawings, observers tended to perceive only one of two basic surface shapes. The interpretations of the drawings may be categorized as elliptic (egg-shaped) and hyperbolic (saddle-shaped) surface patches. These two categories actually belong to a continuum which ranges from umbilical (perfectly spherical) to minimal surfaces (the most symmetrical of all hyperbolic patches) [23]. In between the elliptic and hyperbolic patches lie the parabolic surfaces (i.e. locally cylindrical). The parabolic case constitutes a set of measure zero in the space of all possible local shapes and hence should not be adopted by observers unless evidence is particularly compelling [34]. In the case of line drawings, the latter remark would imply that a parabolic interpretation is taken by the observer only when at least one contour is straight in the image [15].

The categorical perceptions of the line drawings justified the task of our psychophysical experiment. In this experiment, observers were asked to report their first impression for the shape of the surface patch depicted by a given line drawing. The preference to perceive one or the other shape was dependent not only on the geometry of the line drawing (the aspect ratio of the figure), but also on the orientation of the figure in the image plane. The results were not in accord with a world in which the patches were equally likely to occur at any orientation, with any shape and with any surface markings. Instead, subjects displayed consistent biases which compelled them to perceive one surface shape more frequently than another. Thus, the analysis of the frequency of responses provided a window into the biases that observers have concerning the scenes that gave rise to these images.

The biases manifested by the observers were subsequently formalized using a simple Bayesian model. The

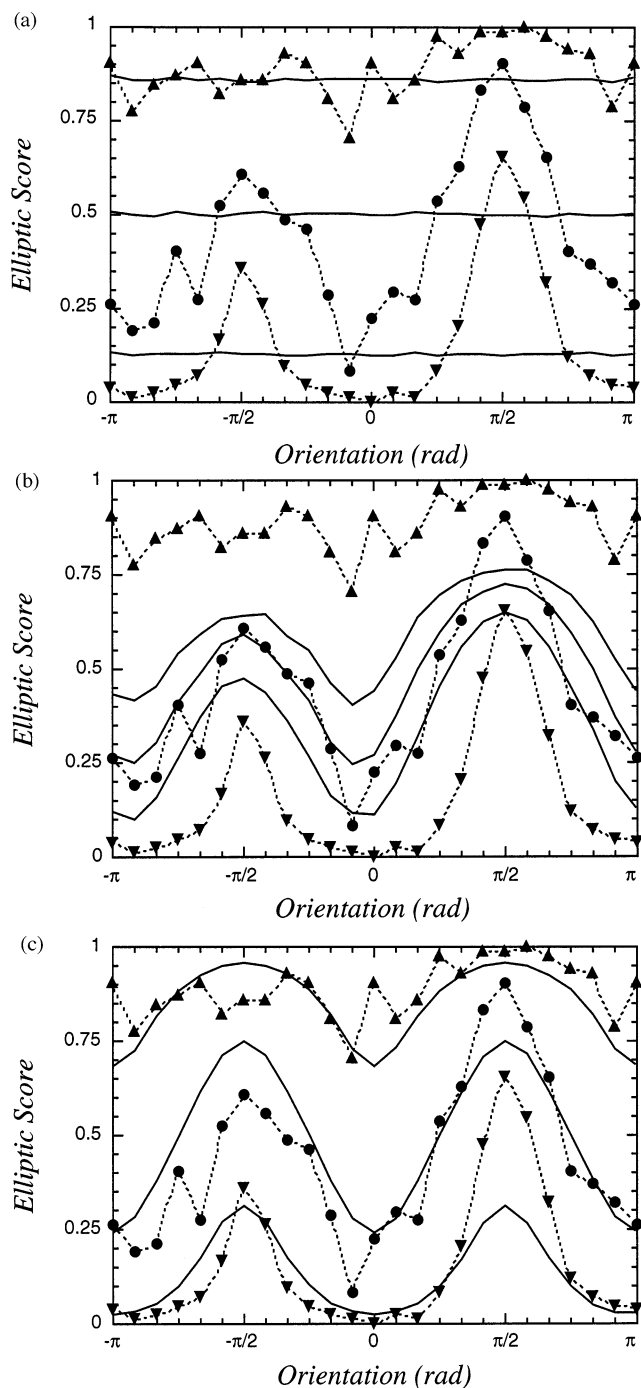


Fig. 13. Simulations of the model with its best fit parameters except one which is set to a value such that the parameter does not have any effect: (a) model with σ_ϵ set to $+\infty$; (b) model with σ_ψ set to $+\infty$; (c) model with μ_k set to 0.0.

model computed the posterior probability that the line drawing corresponded to a surface patch that was elliptic. Due to the simplicity of our images, this endeavor was tractable. This analysis was naturally decomposed into three independent factors: the relationship between the observer and the surface, the local shape of the surface, and the way lines were

drawn on the surface. We modeled potential biases for each of these factors by prior probability functions. The parameters for these functions were then obtained by fitting the model to the observers' data. To complete our model, we had to choose a decision rule that would transform the posterior distribution into a response. We proposed that this rule was noncommitting, i.e. the frequency of observers' responses for ellipticity was identical with the posterior probability that the object was elliptic given the image data. The goodness of fit of the model indicated that all three classes of assumptions participated in the interpretation of the line drawings. These biases were for convex surfaces, markings nearly aligned with the principal curvatures, and for surface normals pointing upwards. Omitting any one of these biases resulted in a substantially poorer fit to the data.

For mathematical convenience, we also assumed that surface contours were geodesics. The geodesic assumption states that the curvature of a surface contour is solely due to the surface curvature. Relaxing the geodesic assumption would be tantamount to allowing a portion of the image curvature to be independent of the surface geometry (a kind of 'curvature noise'), which would reduce the sharpness of the conclusions observers could draw about the underlying qualitative shape. It is possible that the model would still fit as well (e.g. by sharpening the fit values of the local shape prior, i.e. by increasing μ_k). Due to the mathematical complexity involved, it was not possible to simulate a model in which the geodesic constraint was relaxed. It is therefore unclear at this point whether this constraint is really critical for the present model. Nevertheless, it is interesting to note that other empirical evidence suggests that human observers incorporate a geodesic constraint for surface markings [17].

In the current study, the amount of curvature in the image was ignored, and only the sign of curvature figured into the model. This was carried out due to the experimental task using a set of images in which image contour curvature was constant. It would be simple and interesting to elaborate the model to include the image curvature magnitude. The likelihood term would then involve the image contour curvature magnitudes as well as their orientations φ_1 and φ_2 , and the function F would have to match the curvature magnitudes also. This would involve a more elaborate Monte Carlo simulation, but no additional model parameters. As a result, one could sensibly fit new data involving figures with variable image contour curvature.

The three explicit assumptions of our model relate to constraints already discussed in the past. For instance, we interpreted the bias for surface normals to point upwards as a preference to perceive the object as if one were looking at it from above (this interpretation supposes that the general orientation of the object is horizontal, Fig. 5). This constraint could also be related to

the cue of elevation in the picture plane according to which depth usually increases towards the top of an image [35]. The constraint that surface markings approximate lines of curvature was first stated by Stevens [15]. Although surface markings are theoretically unrestricted, Stevens noted a few instances where this constraint was valid, such as wrinkles on the skin or stripes on plants. Finally, the preference for convex surfaces can be related to the salience of 2D objects defined by convex (rather than concave) contours [36–38].

It is natural to wonder how strong these biases are, and how much confidence the visual system puts in these assumptions. While any of them can be easily violated, these assumptions allow us to get a quick first interpretation of the image which can then be revised as more information becomes available. One can also investigate the competition between different assumptions. For instance, we have pitted the surface orientation bias against the well-known bias for illuminant location in shaded images [39]. Numerous studies have reported a preference to perceive a scene as if the light source was located above our head [40–43]. We found that both biases cooperated in the interpretation of ambiguous images containing both contours and shading. For the particular stimuli we used, the viewpoint bias had a stronger effect on the 3D interpretation of the stimuli than the illumination bias. This result suggests that viewpoint-from-above is indeed a robust assumption.

Presumably, shape-from-contour is not the only depth cue that relies on a priori assumptions. Most depth cues provide only restricted information to reconstruct a 3D scene. In image locations where depth cues are sparse, multiple cues cannot interact to fill in missing information, and observers are forced to rely on default assumptions. An attractive hypothesis is that these internal assumptions reflect the statistics of the external world, although it would be difficult to prove. What we have provided in this study is an example of how the human observers' assumptions can be measured.

Acknowledgements

This work was supported by the National Institute of Health grant EY-08266 to M.S. Landy and the Sloan Foundation Theoretical Visual Neuroscience Program. We would like to thank Zili Liu, Bosco Tjan, Laurence Maloney, and two anonymous reviewers for their helpful comments on earlier drafts of the manuscript. The paper contains some results first presented at the European Conference on Visual Perception (ECVP) annual meeting held in September 1995 in Tübingen, Germany [44].

Appendix A. Projection of a curved contour in the image

Our model relies on the likelihood that a stripe painted on a surface projects to a particular contour in the image. This is a problem of projective geometry which can be solved by taking the sequence of rotations that align the tangent plane to the surface with the image. These rotations correspond to the three degrees of freedom that characterize the local orientation of a surface patch, namely the tilt, slant and roll angles (Fig. 7).

We start with a surface contour at orientation ψ relative to the first principal direction (Fig. 9). Assuming that this contour is locally geodesic, its curvature at the origin is simply the normal curvature k of the surface at the orientation ψ . Therefore, as seen from the surface normal, the surface contour appears as a straight line, and its orientation ψ can be restricted to the interval $(-\pi/2, \pi/2)$. Our goal is to compute the orientation ϕ of the projection of this surface contour in the image plane (Fig. 9). We define the orientation ϕ as the angle between the image horizontal and the outward normal to the projected contour. This angle therefore depends on the sign of curvature of the contour, and is defined in the interval $(-\pi, \pi]$.

We define the principal coordinate system (p_1, p_2, N) associated with the tangent plane to the surface patch, where p_1 and p_2 are the principal directions and N is the surface normal. In this coordinate system, a point (x_1, y_1, z_1) on the surface contour near the origin can be written as a function of a parameter t as:

$$\left\{ \begin{array}{l} x_1(t) = \cos(\psi)t \\ y_1(t) = \sin(\psi)t \\ z_1(t) = \frac{1}{2}kt^2. \end{array} \right. \tag{17}$$

The first rotation for the alignment of the tangent plane with the image is the rotation of angle p (the surface roll) about the surface normal. This rotation aligns the first principal direction p_1 with the direction of steepest descent on the surface s_1 (Fig. 14). This latter direction is simply the projection on the surface of the direction of steepest descent in the image i_1 (the projection is taken along the viewing direction). Under this first rotation, the previous coordinate system is transformed into (s_1, s_2, N) , and our point on the surface contour becomes (x_2, y_2, z_2) :

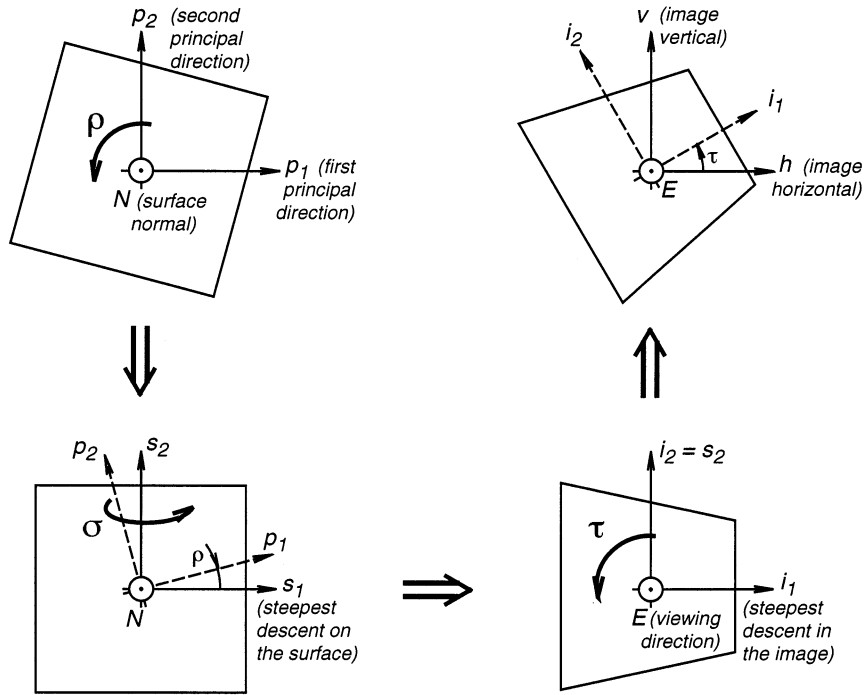


Fig. 14. The function F that relates a given surface marking orientation ψ to an image plane contour orientation ϕ is a concatenation of four rotations, corresponding to the tilt, slant, roll, and ψ itself (not shown here).

$$\begin{cases} x_2(t) = \cos(\psi + \rho)t \\ y_2(t) = \sin(\psi + \rho)t \\ z_2(t) = \frac{1}{2}kt^2. \end{cases} \quad (18)$$

The second step for the alignment procedure is a rotation of angle σ (the surface slant) about the direction s_2 (Fig. 14). In the new coordinate system (i_1, i_2, E) where E is the viewing direction, our point on the surface contour becomes (x_3, y_3, z_3) :

$$\begin{cases} x_3(t) = \cos(\sigma)\cos(\psi + \rho)t + \frac{1}{2}\sin(\sigma)kt^2 \\ y_3(t) = \sin(\psi + \rho)t \\ z_3(t) = -\sin(\sigma)\cos(\psi + \rho)t + \frac{1}{2}\cos(\sigma)kt^2. \end{cases} \quad (19)$$

The final step for the alignment is the rotation of angle τ (the surface tilt) about the viewing direction (Fig. 14). Due to this being a rotation in the image plane, it will simply manifest as an angle τ in the final equation for the orientation of the contour in the image. For simplicity, we shall postpone this last rotation for now, and compute the orientation ϕ' of the contour relative to i_1 .

The orientation of the tangent to the contour is given by the angle θ relative to the direction of steepest descent i_1 (Fig. 15):

$$\theta = \arctan\left(\frac{y_3'(0)}{x_3'(0)}\right) = \arctan\left(\frac{\tan(\psi + \rho)}{\cos(\sigma)}\right). \quad (20)$$

Since the orientation ϕ' is the direction of the outward normal to the contour, ϕ' differs from θ by $\pi/2$. In order to know whether one should add or subtract $\pi/2$, one has to know on which side of the tangent the contour lies. One simple way to obtain this information is to express the contour relative to its tangent, i.e. to

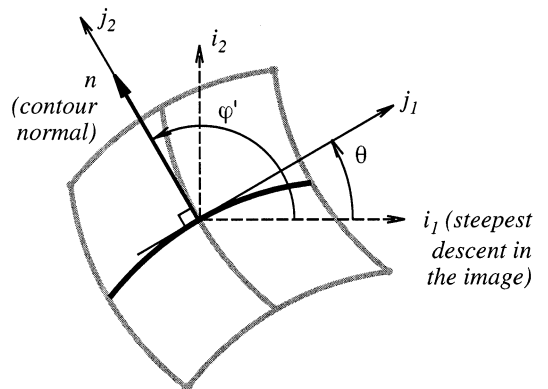


Fig. 15. The orientation ϕ' of the contour relative to the direction of steepest descent in the image i_1 depends on the curvature of the contour. This figure shows the condition when $k \tan(\rho + \psi) \geq 0$, so that ϕ' is $\pi/2$ greater than θ , the direction of the contour tangent.

rotate the coordinate system (i_1, i_2, E) by an angle θ . In this new coordinate system (j_1, j_2, E) , the point (x_4, y_4, z_4) on the contour becomes:

$$\begin{cases} x_4(t) = \cos(\theta)x_3(t) + \sin(\theta)y_3(t) \\ y_4(t) = -\sin(\theta)x_3(t) + \cos(\theta)y_3(t) \\ z_4(t) = z_3(t). \end{cases} \quad (21)$$

We are only interested in the second coordinate (y_4) . Due to the surface slant σ is restricted to $[0, \pi/2]$, we have:

$$\begin{cases} \cos(\theta) = \cos(\sigma) / \sqrt{\cos^2(\sigma) + \tan^2(\psi + \rho)} \\ \sin(\theta) = \tan(\psi + \rho) / \sqrt{\cos^2(\sigma) + \tan^2(\psi + \rho)}, \end{cases} \quad (22)$$

so that we obtain after simplifications:

$$y_4(t) = -\frac{1}{2}k \tan(\psi + \rho)\sin(\sigma) \times t^2 / \sqrt{\cos^2(\sigma) + \tan^2(\psi + \rho)}. \quad (23)$$

The contour will be above its tangent if $y_4(t) \geq 0$, i.e. if $k \tan(\psi + \rho) \leq 0$. Therefore, we have:

$$\varphi' = \arctan\left(\frac{\tan(\psi + \rho)}{\cos(\sigma)}\right) + \text{sign}(k \tan(\psi + \rho))\frac{\pi}{2}. \quad (24)$$

By applying the last rotation to align the tangent plane with the image, we finally obtain the orientation φ of the projected contour relative the image horizontal:

$$\varphi = \arctan\left(\frac{\tan(\psi + \rho)}{\cos(\sigma)}\right) + \text{sign}(k \tan(\psi + \rho))\frac{\pi}{2} + \tau. \quad (25)$$

Appendix B. Monte Carlo simulations

We have formally derived a model to account for the psychophysical experiment. Unfortunately, we did not manage to find a closed form solution for this model. We outline in this appendix the sequence of steps underlying simulations obtained with a Monte Carlo method.

We want to compute the posterior probability that a surface patch is elliptic given image contours oriented as φ_1 and φ_2 . From Bayes' rule, this probability can be written as the following ratio:

$$P(\text{elliptic}|\varphi_1, \varphi_2) = \frac{P(\varphi_1, \varphi_2|\text{elliptic})P(\text{elliptic})}{P(\varphi_1, \varphi_2)}. \quad (26)$$

The numerator is the product of a likelihood term and the prior probability that a random surface patch is elliptic. The likelihood term is the probability that an elliptic patch would produce an image with contours oriented as φ_1 and φ_2 . The denominator is just a normalizing constant so that Eq. (13) is satisfied.

We start with the computation of the likelihood. The idea is to sample the space of elliptic surface patches and to collect the associated orientations of the projected surface contours in the image. Once a large number of surface samples have been drawn, the likelihood will be approximated by the proportion of projected orientations φ_1 and φ_2 .

We begin by picking an elliptic surface patch. We remember that the shape of a patch is fully described by its two principal curvatures k_1 and k_2 . We choose these principal curvatures from a Gaussian distribution with mean μ_k and unit variance (Eq. (5)) with the constraint that the associated shape characteristic χ (Eq. (4)) is positive.

We then paint two contours on this surface patch. Due to the assumption that these contours are locally geodesic (see text), these contours appear as straight lines from the surface normal viewpoint. Therefore, the contours can be simply characterized by their orientation relative to the first principal direction. We choose these orientations ψ_1 and ψ_2 to be Gaussian deviates away from each principal direction, with a standard deviation of σ_ψ (Eq. (6)). From Eq. (12), we can compute the normal curvatures associated with the two contour directions.

The next step is to project the surface patch on the image seen by the observer. This projection involves three rotation angles, starting with the surface roll ρ taken from a uniform distribution over $(0, \pi/2)$. The second angle is the surface slant taken from the distribution given by Eq. (3) (one can easily obtain a sample from this distribution by taking the arc-cosine of a uniform sample in $(0, 1)$). The last angle is the surface tilt taken from the Gaussian distribution with mean $\pi/2$ and standard deviation σ_τ (Eq. (2)). Taking into account these three angles for the surface orientation, one can compute the projected angles φ_1 and φ_2 using Eq. (11).

We repeat the previous steps (pick one surface patch, paint two contours, project the surface patch) over and over again, obtaining many samples of the projected angles φ_1 and φ_2 . These samples are accumulated in a bivariate histogram with bins of width 7.5° , resulting in an estimate of the likelihood of obtaining the angles φ_1 and φ_2 given that the shape is elliptic. In our simulations, we repeated these steps 10 million times.

The prior term in Eq. (26) is the a priori probability that the shape is elliptic. The local solid shape is elliptic if and only if its principal curvatures are both convex or both concave. From Eq. (5), we can deduce the probability $P(\text{convex})$ that one principal curvature is convex (remembering that a contour is convex if its curvature is negative):

$$\begin{aligned}
 P(\text{convex}) &= \int_{-\infty}^0 p(k)dk \\
 &= \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(k-\mu_k)^2}{2}\right)dk \\
 &= \int_{-\infty}^{-\mu_k} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)dt = C(-\mu_k) \\
 &= 1 - C(\mu_k), \tag{27}
 \end{aligned}$$

where $C(x)$ is the cumulative standard normal function. Similarly, we compute that the probability $P(\text{concave})$ that one principal curvature is concave to be $C(\mu_k)$. The prior probability that the surface is elliptic is then:

$$\begin{aligned}
 P(\text{elliptic}) &= P(\text{convex})^2 + P(\text{concave})^2 \\
 &= 1 - 2C(\mu_k) + 2C(\mu_k)^2. \tag{28}
 \end{aligned}$$

References

- [1] Kaufman L. Sight and Mind: An Introduction to Visual Perception. London: Oxford University Press, 1974.
- [2] Marr D. Vision: A computational investigation into the human representation and processing of visual information. San Francisco, CA: WH Freeman, 1982.
- [3] Mamassian P, Kersten D. Illumination, shading and the perception of local orientation. *Vision Res* 1996;36:2351–67.
- [4] Landy MS, Maloney LT, Johnston EB, Young M. Measurement and modeling of depth cue combination: In defence of weak fusion. *Vision Res* 1995;35:389–412.
- [5] Pirenne MHL. Optics, Painting and Photography. Cambridge: Cambridge University Press, 1970.
- [6] Kennedy JM. A Psychology of Picture Perception. San Francisco, CA: Jossey-Bass Publishers, 1974.
- [7] Sedgwick HA, Nicholls AL. Cross talk between the picture surface and the pictured scene: Effects on perceived shape. *Perception* 1993;22(Supplement):109.
- [8] Waltz D. Understanding line drawings of scenes with shadows. In Winston PH, editor. *The Psychology of Computer Vision*. New York: McGraw-Hill, 1975:19–91.
- [9] Barrow HG, Tenenbaum JM. Interpreting line drawings as three-dimensional surfaces. *Artif Intell* 1981;17:75–116.
- [10] Malik J. Interpreting line drawings of curved objects. *Int J Comput Vision* 1987;1:73–103.
- [11] Marr D. Analysis of occluding contour. *Proc R Soc London, B* 1977;197:441–75.
- [12] Koenderink JJ. What does occluding contour tell us about solid shape? *Perception* 1984;13:321–30.
- [13] Witkin A. Recovering surface shape and orientation from texture. *Artificial Intelligence* 1981;17:17–47.
- [14] Stevens KA, Brookes A. Probing depth in monocular images. *Biol Cybern* 1987;56:355–66.
- [15] Stevens KA. The visual interpretation of surface contours. *Artif Intell* 1981;17:47–73.
- [16] Stevens KA. Inferring shape from contours across surfaces. In: Pentland AP, editor. *From Pixels to Predicates*. Norwood, NJ: Ablex, 1986:93–110.
- [17] Knill DC. Perception of surface contours and surface shape: From computation to psychophysics. *J Opt Soc Am A* 1992;9(4):1449–64.
- [18] Nakayama K, Shimojo S. Experiencing and perceiving visual surfaces. *Science* 1992;257(5075):1357–63.
- [19] Freeman WT. The generic viewpoint assumption in a framework for visual perception. *Nature (London)* 1994;368:542–5.
- [20] Kersten D. Statistical limits to image understanding. In: Blake-More C, editor. *Vision: Coding and Efficiency*. Cambridge: Cambridge University Press, 1990:32–44.
- [21] Kersten D. Transparency and the cooperative computation of scene attributes. In: Landy MS, Movshon JA, editors. *Computational Models of Visual Processing*. Cambridge, MA: MIT Press, 1991:209–228.
- [22] Knill DC, Richards W. *Perception as Bayesian Inference*. Cambridge: Cambridge University Press, 1996.
- [23] Mamassian P, Kersten D, Knill DC. Categorical local-shape perception. *Perception* 1996;25:95–107.
- [24] Cohen JD, MacWhinney B, Flatt M, Provost J. PsyScope: A new graphic interactive environment for designing psychology experiments. *Behav Res Methods, Instrum Comput* 1993;25:257–71.
- [25] Knill DC, Kersten D, Yuille A. Introduction: A Bayesian formulation of visual perception. In: Knill DC, Richards W, editors. *Perception as Bayesian Inference*. Cambridge: Cambridge University Press, 1996:1–21.
- [26] Yuille AL, Bülthoff HH. Bayesian decision theory and psychophysics. In: Knill DC, Richards W, editors. *Perception as Bayesian Inference*. Cambridge: Cambridge University Press, 1996:123–161.
- [27] Mamassian P, Landy MS. Illuminant and viewpoint biases from embossed surfaces. *European Conference on Visual Perception. Abstr Perception* 1997;26(Supplement):51.
- [28] Batschelet E. *Circular Statistics in Biology*. London: Academic Press, 1981.
- [29] Hilbert D, Cohn-Vossen S. *Anschauliche Geometrie*. Berlin: Springer (1932). English translation: *Geometry and the Imagination*, New York: Chelsea, 1952.
- [30] Koenderink JJ. *Solid Shape*. Cambridge, MA: MIT Press, 1990.
- [31] do Carmo MP. *Differential Geometry of Curves and Surfaces*. Englewood Cliffs, NJ: Prentice-Hall, 1976.
- [32] Berger JO. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer, 1985.
- [33] Mood AM, Graybill FA, Boes DC. *Introduction to the theory of statistics*. New York: McGraw Hill, 1974.
- [34] Bennett BM, Hoffman DD, Prakash C. *Observer Mechanics: A Formal Theory of Perception*. New York: Academic Press, 1989.
- [35] Berbaum K, Tharp D, Mroczek K. Depth perception of surfaces in pictures: Looking for conventions of depiction in Pandora's box. *Perception* 1983;12:5–20.
- [36] Rubin E. *Visuell wahrgenommene Figuren: Studien in psychologischer Analyse*. Kobenhavn (Copenhagen): Gyldendalske Boghandel, 1921.
- [37] Kanizsa, G., Gerbino, W. Convexity and symmetry in figure-ground organization. In: Henle M, editor. *Vision and Artifact*. New York: Springer, 1976:25–32.
- [38] Liu Z, Jacobs DW, Basri R. Perceptual completion: Beyond good continuation. *ARVO annual meeting. Invest Ophthalmol Visual Sci* 1995;36(Supplement):475.
- [39] Mamassian P, Landy MS. Cooperation of priors for the perception of shaded line drawings. *European Conference on Visual Perception. Abstr Perception* 1996;25(Supplement):21.
- [40] Rittenhouse D. Explanation of an optical deception. *Trans Am Philos Soc* 1798;2:37–42.
- [41] Brewster D. On the optical illusion of the conversion of cameos into intaglios and of intaglios into cameos with an account of other analogous phenomena. *Edinburgh J Sci* 1826;4:99–108.
- [42] Gibson JJ. *The Perception of the Visual World*. Boston, MA: Houghton Mifflin, 1950.
- [43] Ramachandran VS. Perception of shape from shading. *Nature (London)* 1988;331:163–6.
- [44] Mamassian P. Solid shape from occluding contour. *European Conference on Visual Perception. Abstr Perception* 1995;24(Supplement):35.