MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

The Generic Viewpoint Assumption in a Bayesian Framework

William T. Freeman

TR93-11a December 1994

Abstract

The g̈eneric viewässumption states that an observer is not in a special position relative to the scene. It is commonly used to disqualify scene interpretations that assume special viewpoints, following a binary decision that the viewpoint was either generic or accidental. In this chapter, we show how to use the generic view assumption to quantify the likelihood of a view. This quantitative approach can be applied to estimate scene parameters. This approach applies to many vision problems. We show shape from shading examples where we rank shapes or reflectance functions in cases which are otherwise ambiguous. The rankings agree with the perceived values.

Bayesian Perspectives on Perception, David Knill and Whitman Richards, editors, Cambridge University Press, 1996

Publication History:–

1. First printing, TR-93-11, June 1993
2. Second printing, TR-93-11a, March 1994

Contents

1 Introduction

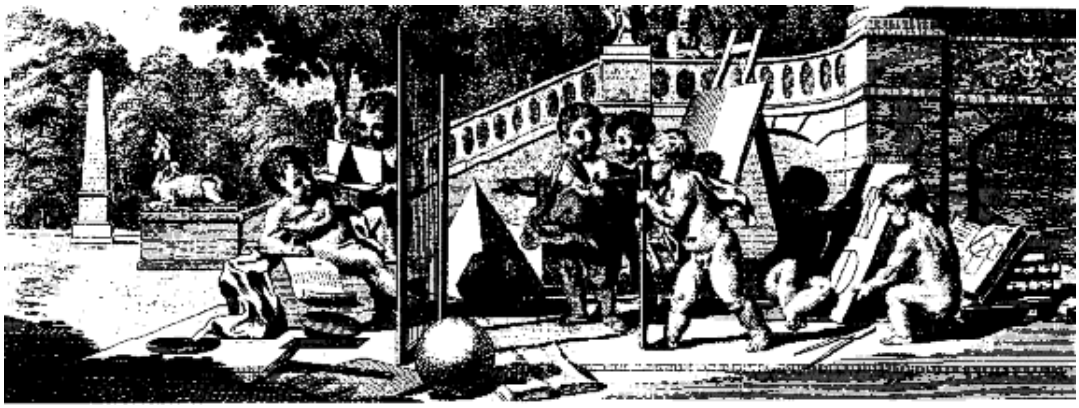


Figure 1: This image can describe one approach to the problem of perception. Starting from the visual image, held by the cherubs at the right, one wants to find the probability that each of various models or scene interpretations (on the ground in the middle) could have generated that visual data. (Etching by Samuel Wale, *Putti engaged in the study of geometry and perspective*, from J. Kirby's *The Perspective of Architecture*, London, 1761.)

A task of visual perception is to find the scene which best explains visual observations. Fig. 1 can be used to illustrate the problem of perception. The visual data is the image held by two cherubs at the right. Scattered in the middle are various geometrical objects–"scene interpretations"–which may account for the observed data. How does one choose between the competing interpretations for the image data?

One approach is to find the probability that each interpretation could have created the observed data. Bayesian statistics are a powerful tool for this, e.g. [17, 46, 30, 28]. One expresses prior assumptions as probabilities and calculates for each interpretation a posterior probability, conditioned on the visual data. The best interpretation may be that with the highest probability density, or a more sophisticated criterion may be used. Other computational techniques, such as regularization [47, 42], can be posed in a Bayesian framework [46]. In this chapter, we will apply the powerful assumption of "generic view" in a Bayesian framework. This will lead us to an additional term from Bayesian theory involving the Fisher information matrix. This will modify

the posterior probabilities to give additional information about the scene.

The generic view assumption [31, 4, 3, 36, 43, 38, 39] postulates that the scene is not viewed from a special position. Fig. 2 shows an example. The square in (a) could be an image of a wire-frame cube (b) viewed from a position where the line segments of the front face hide those behind them. However, that would require a very special viewpoint, and given the image in (a), one should infer a square, not a cube; references [21, 39] discuss this example. The generic view assumption has been invoked to explain perceptions involving stereo and transparency [39], linear shape from shading [41], and feature or object identification [31, 36, 3, 43, 38, 28, 12]. Typically, researchers assume a view is either generic, and therefore admissible, or accidental, and therefore rejected. Some have pointed out that it should be possible to quantify the degree of accidentalness or have done so in special cases [49, 36, 38, 34, 39, 28, 12].

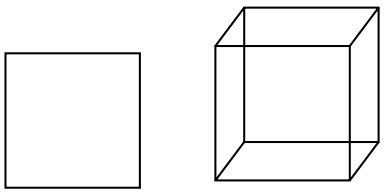


Figure 2: An example of use of the generic view assumption for binary decisions. The image (a) could be of a square, or it could be an "accidental view" of the cube in (b). Since a cube would require a special viewing position to be seen as the image in (a), we reject that possible interpretation for (a).

In this chapter we take a quantitative approach, in a complementary fashion to the categorical use of the generic viewpoint assumption. Rather than assign categories to views of geometrical objects, we assign probabilities to views of continuous surfaces. This approach will let us make parametric decisions about candidate scene interpretations. We discuss the relationship to aspect graphs in Section 4.

In our approach, we divide parameters into two groups, *generic* parameters and *scene* parameters. Generic parameters are parameters we do not need to estimate precisely. Viewpoint is one example; others are object orientation and lighting position. Scene parameters are what we desire to

estimate, and can be reflectance function, shape, lighting direction and velocity. We relax this split into two types of parameters in the formulation of Section 6.

We will exploit the known distributions of the generic variables to modify the probability of given scene parameters. We will used an established approximation from Bayesian statistics to quantify the "genericity" of a view. Our analysis will favor scene interpretations for which the visual data is stable to small changes in the generic variables. This information can be used to make principled selections among competing scene interpretations where otherwise arbitrary choices would have to be made.

We show applications to the shape from shading problem. Using a two-parameter family of reflectance functions, we show how to find the probability of a shape and reflectance function combination from a single image. We rank shape probabilities corresponding to different assumed lighting directions in a case where each shape can account for the image data equally well.

We motivate our approach in the remainder of the introduction. In Section 2 we derive the *scene probability equation*, the conditional probability for a scene interpretation given the observed visual data. Then we show the applications to shape from shading. In Section 6, we frame this approach in terms of the loss functions of Bayesian estimation. Substantial parts of this work were presented in [15, 16].

1.1 Example

A simple example illustrates the main idea. Suppose the visual data is the image of Fig. 3 (a). Perceptually, there are two possible interpretations: it could be a bump, lit from the left, or a dimple, lit from the right. Yet mathematically, there are many interpretations to choose from under the commonly encountered conditions of linear shading [1] The image could arise from any of the shapes shown in (b), under the proper lighting conditions, which are indicated by the lighting direction arrow shown next to each shape. How should one choose between these competing explanations?

One often considers only two criteria to evaluate an interpretation: how well it accounts for the observed data, and the prior probability that the

[1] In linear shading [41], the image intensity is a linear function of the local surface gradient. For small surface slopes and low angles of illumination, linear shading is a good approximation to Lambertian shading.

interpretation would exist in the world. If each shape accounts equally well for the image data, we are left with choosing based on the prior probabilities for each shape. We could arbitrarily decide that we like bump shapes more than tube shapes but we have no grounds for that. Such an arbitrary decision could easily lead to an incorrect interpretation for some other image. What is missing from this approach?

For the three tube-like shapes shown, there is a suspicious alignment between the inferred surface structure and the assumed light direction. This seems unlikely, and we would like to include this coincidence in our probability calculation. Fig. 3 (c) and (d) give an intuition for how we might measure the accidentalness of the surface and light direction alignments. If we imagine wiggling the assumed light direction slightly, we see that for the shape of (c), the image changes quite a bit. For the shape of (d), we can observe the image of (a) over a much broader range of assumed light directions. If the light had an equal chance of coming from each different direction then there are more opportunities for the shape of (d) to have presented us with the image (a) than there are for the shape of (c).

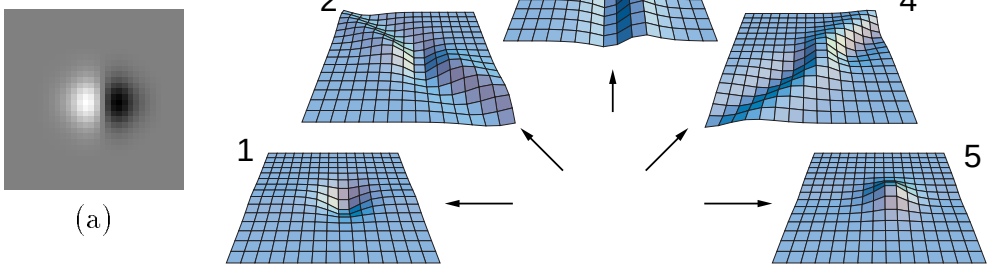
We will quantify this intuition by taking derivatives. A small image intensity derivative with respect to light direction means that the image will look almost the same over a relatively large range of light angles. If all light angles are equally likely, then the likelihood of a shape is proportional to the range of light angles over which the shape looks nearly the same as the image data. In our analysis, the image derivatives will result from writing the image in a Taylor series in the generic variable.

In this way, we will exploit additional assumptions about the visual world. To the prior assumptions about what is being estimated, we add assumptions about relationships between the object, the viewer and the light source direction. This additional information may allow for weakening the prior assumptions we need to make about the scene.

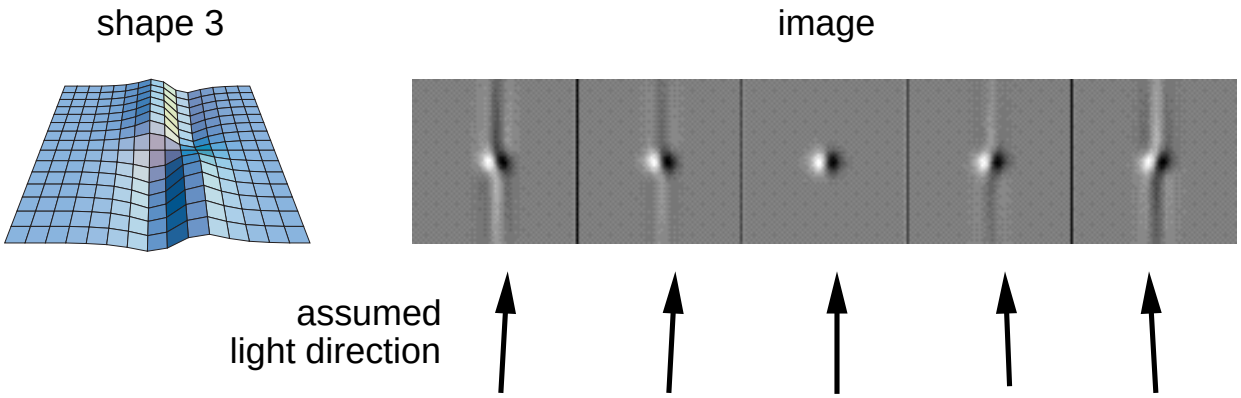
This approach, and the approximation we use, comes from the literature of Bayesian statistics. Using the Taylor series approximates the likelihood term of the posterior probability as a gaussian, which gives a Bayesian version of the central limit theorem. This was done by Laplace ([33], cited in [1]), in the 1800's as well as Fisher [13] and Jeffreys [26] in the first half of this century. Others have followed and extended this approach [7, 35, 29, 1, 20, 45, 37] (Berger [1] and MacKay [37] are particularly accessible references).

In the field of computer vision, Szeliski [46] applied maximum likelihood

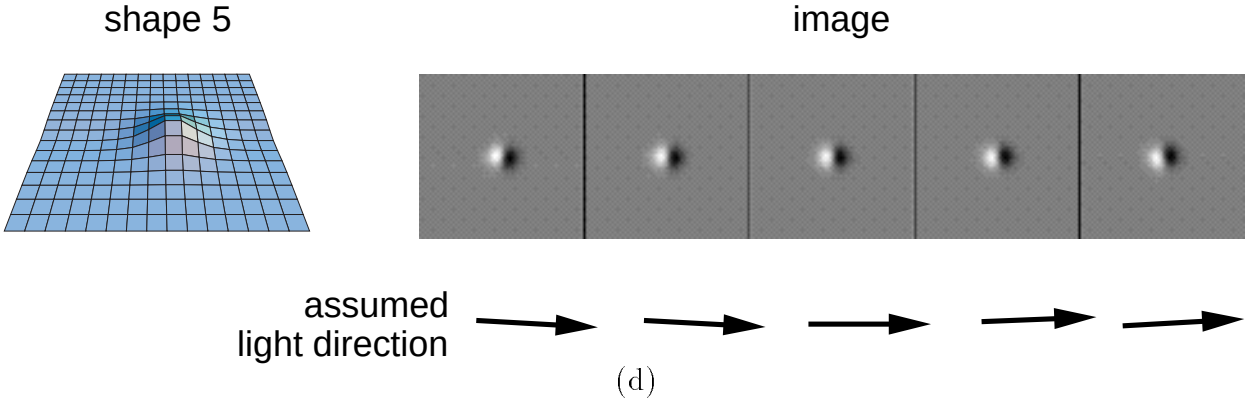
estimation to favoring interpolation parameters which could have generated the observed data in many different ways. See Weinshall *et. al.* [48] for a recent related non-Bayesian approach.


3
2
4
1
5
shapes for different assumed light directions

(a) (b)


shape 3
image
assumed
light direction

(c)


shape 5
image
assumed
light direction

(d)

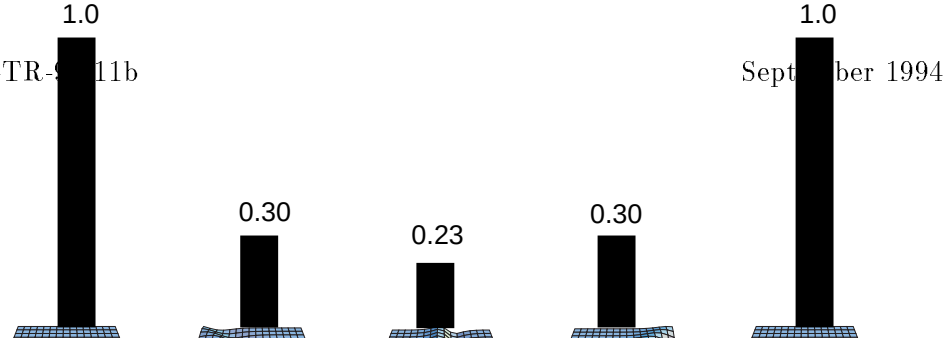

1.0
0.30
0.23
0.30
1.0

Figure 3: (a) Perceptually, this image has two possible interpretations: a bump lit from the left, or a dimple lit from the right. But under the conditions of linear shading [41], each of the shapes shown in (b), under shallow illumination from the direction indicated, could create the image (a). (The boundary conditions used were those described in [41]). Shapes $2-4$ require coincidental alignment with the assumed light direction. We can distinguish between the explanations if we examine the image stability to perturbations in light direction. Shape 3 presents the observed image for a very small range of light angles (c), while shape 5 does for a much larger range (d). The scene probability equation, which we derive in Section 2, quantifies the degree of coincidence of the light direction and reconstructed shape. Using light direction as the generic variable, and shape as the scene parameter, the scene probability for each interpretation is plotted in (e). Note the preference for the bump or dimple shapes, in agreement with the appearance of (a).

2 The Scene Probability Equation

In this section we derive the probability densities for scene parameters given observed data. We describe the problem generally.

Let $\hat{\vec{y}}$ be a vector of observations. This can be image intensities, or measures derived from them, such as spatial or temporal derivatives or normal velocities. For simplicity, we will often call this "the image".

Let the vector $\vec{\beta}$ be the scene parameters we want to estimate. This vector can describe, for example, the object shape and reflectance function or the image velocities.

Let $\vec{x}$ be a vector of the generic variables–variables which we do not need to estimate precisely. For the example of Fig. 3 this was the incident light angle. Generic variables can include viewpoint position, object orientation, lighting position, or texture orientation. For now we assume that the probability density of the generic variables $\vec{x}$, $P_{\vec{x}}(\vec{x})$, is flat:

$$P_{\vec{x}}(\vec{x}) = k \tag{1}$$

where k is a normalization constant. (The notation $P_a(x)$ denotes the probability density of the variable a as a function of x. For brevity, we omit the subscript for conditional probability functions.) We indicate later how to generalize to the case of non-flat generic variable densities.

The scene parameters $\vec{\beta}$ and generic variables $\vec{x}$ determine the ideal observation (image), $\vec{y}$, through a "rendering function", $\vec{f}$:

$$\vec{y} = \vec{f}(\vec{x}, \vec{\beta}) \tag{2}$$

For the example of Fig. 3 the rendering function was the computer graphics calculation which gave the image as a function of surface shape, $\vec{\beta}$, and incident light angle, $\vec{x}$.

We postulate some measurement noise, although we will often examine the limit where its variance goes to zero. The observation, $\hat{\vec{y}}$, is the rendered ideal image $\vec{y}$ plus the measurement noise, $\vec{n}$:

$$\hat{\vec{y}} = \vec{y} + \vec{n}. \tag{3}$$

The noise specifies a distance metric between images–the probability that the differences between the images are due to noise. We will assume that the

measurement noise is a set of independent, identically distributed Gaussian random variables with mean zero and standard deviation σ. Thus $P_{\vec{n}}(\vec{n})$, the probability density function of the noise, is

$$P_{\vec{n}}(\vec{n}) = \frac{1}{(\sqrt{2\pi\sigma^2})^N} \exp \frac{-\|\vec{n}\|^2}{2\sigma^2}, \tag{4}$$

where N is the dimension of the observation and noise vectors and $\|\vec{n}\|^2 = \vec{n}\cdot\vec{n}$.

Given an observation $\hat{\vec{y}}$, we want to find $P(\vec{\beta} \mid \hat{\vec{y}})$, the conditional probability of the parameters $\vec{\beta}$. We first use Bayes' theorem to evaluate the joint probability of $\vec{\beta}$ and a particular value of the generic variables, $P(\vec{\beta}, \vec{x} \mid \vec{y})$:

$$P(\vec{\beta}, \vec{x} \mid \vec{y}) = \frac{P(\vec{y} \mid \vec{\beta}, \vec{x}) P_{\vec{\beta}}(\vec{\beta}) P_{\vec{x}}(\vec{x})}{P_{\vec{y}}(\vec{y})}, \tag{5}$$

where we have assumed that $\vec{x}$ and $\vec{\beta}$ are independent. The denominator is constant for all models $\vec{\beta}$ to be compared.

To find $P(\vec{\beta}, \vec{x} \mid \vec{y})$, independent of the value of the generic variable $\vec{x}$, we integrate the joint probability of Eq. (5) over the possible $\vec{x}$ values:

$$P(\vec{\beta}|\vec{y}) = \frac{P_{\vec{\beta}}(\vec{\beta})}{P_{\vec{y}}(\vec{y})} \int P(\vec{y} \mid \vec{\beta}, \vec{x})\, P_{\vec{x}}(\vec{x})\, d\vec{x}. \tag{6}$$

From our Gaussian noise model, Eq. (4), we have

$$P(\vec{y} \mid \vec{\beta}, \vec{x}) = \frac{1}{(\sqrt{2\pi\sigma^2})^N} e^{\frac{-\|\vec{y} - \vec{f}(\vec{x}, \vec{\beta})\|^2}{2\sigma^2}}. \tag{7}$$

$P(\vec{y} \mid \vec{\beta}, \vec{x})$ is large where the scene $\vec{\beta}$ and the value $\vec{x}$ yield a rendered image similiar to the observation $\vec{y}$. The integral of Eq. (6) measures the area of $\vec{x}$ for which the scene $\vec{\beta}$ yields the observation. In our example, it effectively counts the frames in Figure 3 (c) or (d) where the rendered image is similar to the input data. This will favor shape 5, for which the image changes little over a range of light angles.

For the low noise limit, we can find an analytic approximation to the integral of Eq. 6 [27, 5]. We expand $\vec{f}(\vec{x}, \vec{\beta})$ in Eq. (7) in a second order

Taylor series,

$$\vec{f}(\vec{x},\vec{\beta}) \approx \vec{f}(\vec{x}_0,\vec{\beta}) + \sum_i \vec{f'}_i\,[\vec{x}-\vec{x}_0]_i + \frac{1}{2}\sum_{i,j}[\vec{x}-\vec{x}_0]_i\,\vec{f''}_{ij}\,[\vec{x}-\vec{x}_0]_j, \qquad (8)$$

where $[\cdot]_i$ indicates the ith component of the vector in brackets, and

$$\vec{f'}_i = \frac{\partial \vec{f}(\vec{x},\vec{\beta})}{\partial x_i}\,|_{\vec{x}=\vec{x}_0}, \qquad (9)$$

and

$$\vec{f''}_{ij} = \frac{\partial^2 \vec{f}(\vec{x},\vec{\beta})}{\partial x_i \partial x_j}\,|_{\vec{x}=\vec{x}_0}\ . \qquad (10)$$

We take $\vec{x}_0$ to be the value of $\vec{x}$ which can best account for the observed image data; i.e., for which $\|\vec{y}-\vec{f}(\vec{x},\vec{\beta})\|^2$ is minimized.

Using Eqs. (7)–(10) to second order in $\vec{x}-\vec{x}_0$ in the integral of Eq. (6), we find the posterior probability for the scene parameters $\vec{\beta}$ given the visual data $\vec{y}$:

$$\begin{aligned} P(\vec{\beta}\mid\vec{y}) &= k\,\exp\left(\frac{-\|\vec{y}-\vec{f}(\vec{x}_0,\vec{\beta})\|^2}{2\sigma^2}\right) P_{\vec{\beta}}(\vec{\beta})\,\frac{1}{\sqrt{\det(\boldsymbol{C})}} \qquad (11)\\ &= k\,(\text{fidelity})\,(\text{prior probability})\,(\text{generic view}), \end{aligned}$$

where the i and jth elements of the matrix $\boldsymbol{C}$ are

$$C_{ij} = \vec{f'}_i \cdot \vec{f'}_j - (\vec{y}-\vec{f}(\vec{x}_0,\vec{\beta}))\cdot\vec{f''}_{ij}. \qquad (12)$$

We call this the *scene probability equation*. We have combined the constants which do not depend on the scene $\vec{\beta}$ into the normalization constant k, which can be set so that the integral of $P(\vec{\beta}\mid\hat{\vec{y}})$ over all $\vec{\beta}$ is one. Usually we examine relative probabilities; then k doesn't matter. The rendering function derivatives in Eq. (8) must exist for Eq. (11) to hold. Thus, we cannot apply Eq. (11) to some idealized geometrical objects.

The scene probability equation Eq. (11) has two familiar terms and a new term in computer vision. The term $\exp(\frac{-\|\hat{\vec{y}}-\vec{f}(\vec{x}_0,\vec{\beta})\|^2}{2\sigma^2})$ penalizes scene hypotheses which do not account well for the original data (hypotheses $\vec{\beta}$ for which the squared difference of $\vec{f}(\vec{x}_0,\vec{\beta})$ from the image data $\hat{\vec{y}}$ is large). We

call this the *image fidelity term.* (This may also be called the "likelihood of $\vec{x}_0$ and $\vec{\beta}$ with respect to $\hat{\vec{y}}$"). The *prior probability* term $P_{\vec{\beta}}(\vec{\beta})$ came from Bayes' law and incorporates prior assumptions. These two terms (the prior and a squared error term) are familiar in computer vision. $\frac{1}{\sqrt{\det(\boldsymbol{C})}}$ is the new term, arising from the generic view assumption. If the rendered image changes quickly with the generic view variables, then $\frac{1}{\sqrt{\det(\boldsymbol{C})}}$ will be large and the scene hypothesis $\vec{\beta}$ will be unlikely. This $\frac{1}{\sqrt{\det(\boldsymbol{C})}}$ term quantifies our intuitive notion of generic view, and we call it the *generic view* term. Note that the generic view term depends on the best value of the generic variable, $\vec{x}_0$, as well as the scene parameter $\vec{\beta}$, so it is not equivalent to a prior on the scene parameter, which will only be a function of $\vec{\beta}$. The scene probability equation gives the probability that a scene interpretation $\vec{\beta}$ generated the visual data, $\hat{\vec{y}}$, based on fidelity to the data, prior probability, and the probability that the scene would have presented us with the observed visual data.

This quantification of the genericity of a view follows established techniques in Bayesian statistics. The matrix $\boldsymbol{C}$ is called the conditional Fisher information matrix [13, 1]. It is used to approximate the likelihood locally as a Gaussian [13, 26, 7] and can be used in integration over a loss function or in marginalization [35, 1]. For example, Box and Tiao [6] employ this approximation when they integrate out nuisance parameters from a joint posterior, as we have done here. Gull [19] calls $\frac{1}{\sqrt{\det(\boldsymbol{C})}}$ the Occam factor and he, Skilling [45], and MacKay [37] use it as we have here and in other ways. This factor also arises in the context of "non-informative priors" [26, 7, 1].

The case of only one generic variable and $\|\hat{\vec{y}} - \vec{f}(x_0, \vec{\beta})\| = 0$ shows the role of the image derivatives more clearly. Then the scene probability equation becomes:

$$P(\vec{\beta} \mid \hat{\vec{y}}) = k\, P_{\vec{\beta}}(\vec{\beta})\, \frac{1}{\sqrt{\sum_i (\frac{\partial f_i(x,\vec{\beta})}{\partial x} \mid_{x=x_0})^2}}, \tag{13}$$

The probability of a parameter vector $\vec{\beta}$ varies inversely with the sum of the squares of the image derivatives with respect to the generic variable.

In evaluating the Gaussian integral in Eq. (6) we assumed that the generic variables were separable. Generic object pose in 3-d is an exception to that and we derived the scene probability equation for that case in [14]. If the

prior probability of the generic variable were not constant then the factor $P_{\vec{x}}(\vec{x_0})$ would be included in the prior term of Eq. (11).

After finding $P(\vec{\beta} \mid \hat{\vec{y}})$, one can estimate a "best" value $\vec{\beta}$, after making assumptions about the loss incurred by an incorrect estimate (see Section 6). Alternatively, one can pass a representation of the entire probability density function $P(\vec{\beta} \mid \hat{\vec{y}})$ on to a higher level of processing.

3 Shape from Shading Examples

We apply the scene probability equation to some problems in shape from shading. Given a shaded image, lighting conditions and the reflectance function, there are many algorithms which can compute a shape to account for the shaded image; see [22, 23] for reviews.

Most shape from shading algorithms require specification of the lighting and object surface characteristics. There are a number of methods that can infer these given more than one view of the object or other information [25, 50, 18, 40, 52]. Finding the object shape from a single view without these parameters is not a solved problem. Brooks and Horn [10] proposed a more general scheme that iterated to find a shape and reflectance map that could account for the image data.

However, accounting for image data is not enough. For some classes of images, many shapes and reflectance functions can account equally well for an image (although some images which are impossible to explain by Lambertian shading have been found, [24, 9]). An infinite number of surface and light source combinations can explain regions of 1-dimensional intensity variations, since the solution just involves a 1-dimensional integration. The rendering conditions of "linear shading" [41] can be invoked to explain any image, as we discuss later. Thus, to explain a given image, one must choose between a variety of feasible surface shapes, reflectance functions and lighting conditions.

To make such choices, one could invoke prior preferences for the preferred shapes or reflectance functions. Some shape from shading algorithms do this implicitly by using regularizing functionals. However, if one relies too heavily on the prior statistics to make decisions, that will bias the shape reconstructions. The scene probability equation enables one to use the generic view assumption to choose between shapes and reflectance functions, lessening the reliance on the priors.

We have not developed a shape from shading algorithm which uses the scene probability equation directly. Rather, we use existing shape from shading algorithms ([2, 41]) to generate hypothesis shapes and use the scene probability equation to evaluate their probability density. Future research can incorporate the scene probability equation, or an approximation to it, directly into a shape from shading algorithm.

3.1 Reflectance Function

Fig. 4 shows an example of a 1-d image which can be explained by several different shapes and reflectance functions. The image (a) may look like a cylinder (c) painted with a Lambertian reflectance function (b) (shown on a hemisphere). However, it could also have been created by the flatter shape of (f), painted with a shiny reflectance function (e). If both interpretations account for the data, how can we choose between them? We could invoke prior assumptions about the probabilities of various shapes or reflectance functions, but we would prefer a choice based on the image, not our prior assumptions (see [39]).

Before applying the scene probability equation to this example, we provide intuitive motivation for the result. We can distinguish between the two scene hypotheses if we imagine rotating them. The Lambertian shaded image would change little for small rotations, while the shiny image would change considerably, Fig. 4 (d) and (g). Thus, for the Lambertian solution, for a large range of object poses we would see the image of Fig. 4 (a). For the shiny solution, we would see that image over a smaller range of poses.

One might think that the images of rotated shiny objects would always change more than those of Lambertian ones, since the diffuse reflection changes slowly with surface orientation. However, Fig. 5 shows that this is not the case. The image data, Fig. 5 (a), may look like a shiny cylinder, but, again, it can be explained by either a Lambertian reflectance function, shape (c) painted with (b), or a shiny one, the shape (f) painted with (e). Note that the shape for the Lambertian function is taller than that of the shiny reflectance function. When we rotate both shapes, in Fig. 5 (d) and (g), it is the image of the Lambertian shape, (c), which changes more than that of the shiny one (f), because of the parallax induced as the tall shape moves back and forth.

To quantify these intuitions, we can apply the scene probability equation to distinguish between these shapes and reflectance functions. Our observation $\vec{\hat{y}}$ is the image data. The parameter $\vec{\beta}$ we wish to estimate is the shape and reflectance function of the object. We use a two variable parameterization of reflectance functions, a subset of the Cook and Torrance model [11]. The parameters are surface roughness, which governs the width of the specular highlight, and specularity, which determines the ratio of the diffuse and specular reflections. Fig. 6 gives a visual key.

We want to use the scene probability equation, Eq. (11), to evaluate the probability $P(\vec{\beta} \mid \hat{\vec{y}})$ for each reflectance function in our parameterized space. A shape exists for each reflectance function which could have created the 1-d images of Fig. 4 (a) and Fig. 5 (a). (For each shape we assumed boundary conditions of constant height at the vertical picture edge). We will consider only shape and reflectance function combinations which exactly account for the image data. Then the fidelity term in Eq. (13) for $P(\vec{\beta} \mid \hat{\vec{y}})$ is 1. For this example, we will assume a flat prior for the reflectance functions and shapes as parameterized, $P_{\vec{\beta}}(\vec{\beta}) = c$. We used a shape from shading algorithm [2] to find the shape corresponding to each reflectance function.

Now we consider the generic view term. We will use both the vertical rotation of the object and the vertical light position as the generic variables. We need the derivative of the image intensities, I, with respect to the rotation angle, ϕ. This derivative is a special case of a formula given in [14],

$$\frac{dI}{d\phi} = \frac{\partial I}{\partial Y} Z + \frac{\partial m}{\partial q}(1 + q^2), \tag{14}$$

where $q = \frac{\partial Z}{\partial Y}$, X and Y are Cartesian image coordinates, Z is the surface height, and m is the reflectance map. We calculated numerically the image derivative with respect to light position. For the z value of the axis of rotation we used the value which minimized the squared derivative of the image with respect to object rotation angle.

Using the above in the scene probability equation, we plot in Fig. 7 (b) and (d) the probability that each reflectance function and corresponding shape generated the 1-d images, shown again in Fig. 7 (a) and (c). Note that for each image, the high probabilities correspond to reflectance functions which look (see Fig. 6) more like the material of the image patches in Fig. 7 (a) and (c). We have successfully evaluated the relative probabilities that different reflectance functions and shapes created a given image. Note this was done from a single view and for a case where the reflectance function is otherwise completely unknown.

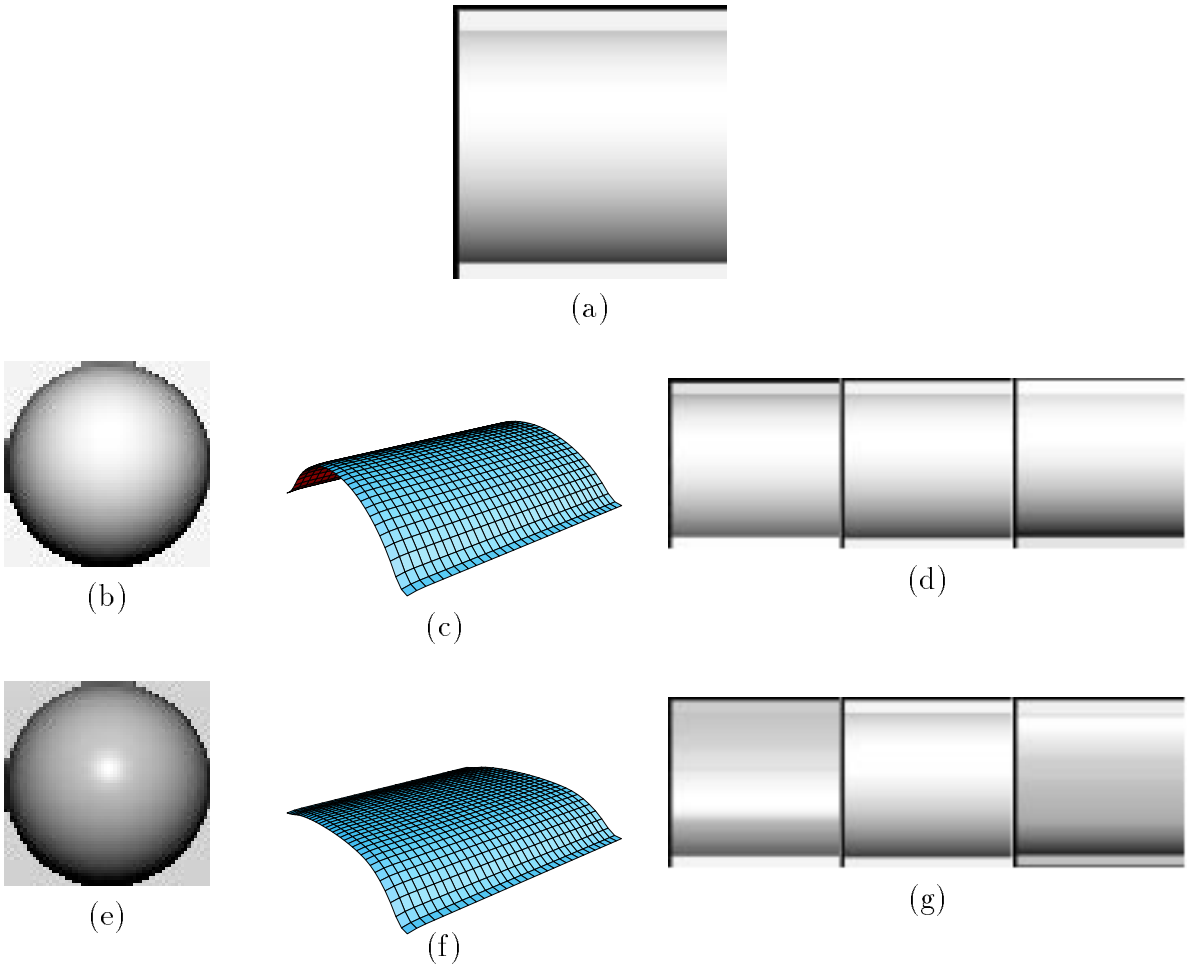


Figure 4: The image (a) may appear to be a cylinder (c) painted with a Lambertian reflectance function (b) (shown on a hemisphere). However the flatter shape of (f) and a shiny reflectance function (e) also explain the data equally well. We can distinguish between the competing accounts for (a) by imagining rotating each shape. Images of each shape at three nearby orientations are shown in (d) and (g). We see that the image made assuming a Lambertian reflectance function (b) is more stable than that made assuming a shiny reflectance function (e). If all object angles are equally likely, and the shapes and reflectances of (c) and (f) are equally likely to occur in the world, then (c) should be a more likely interpretation of (a) than (f).

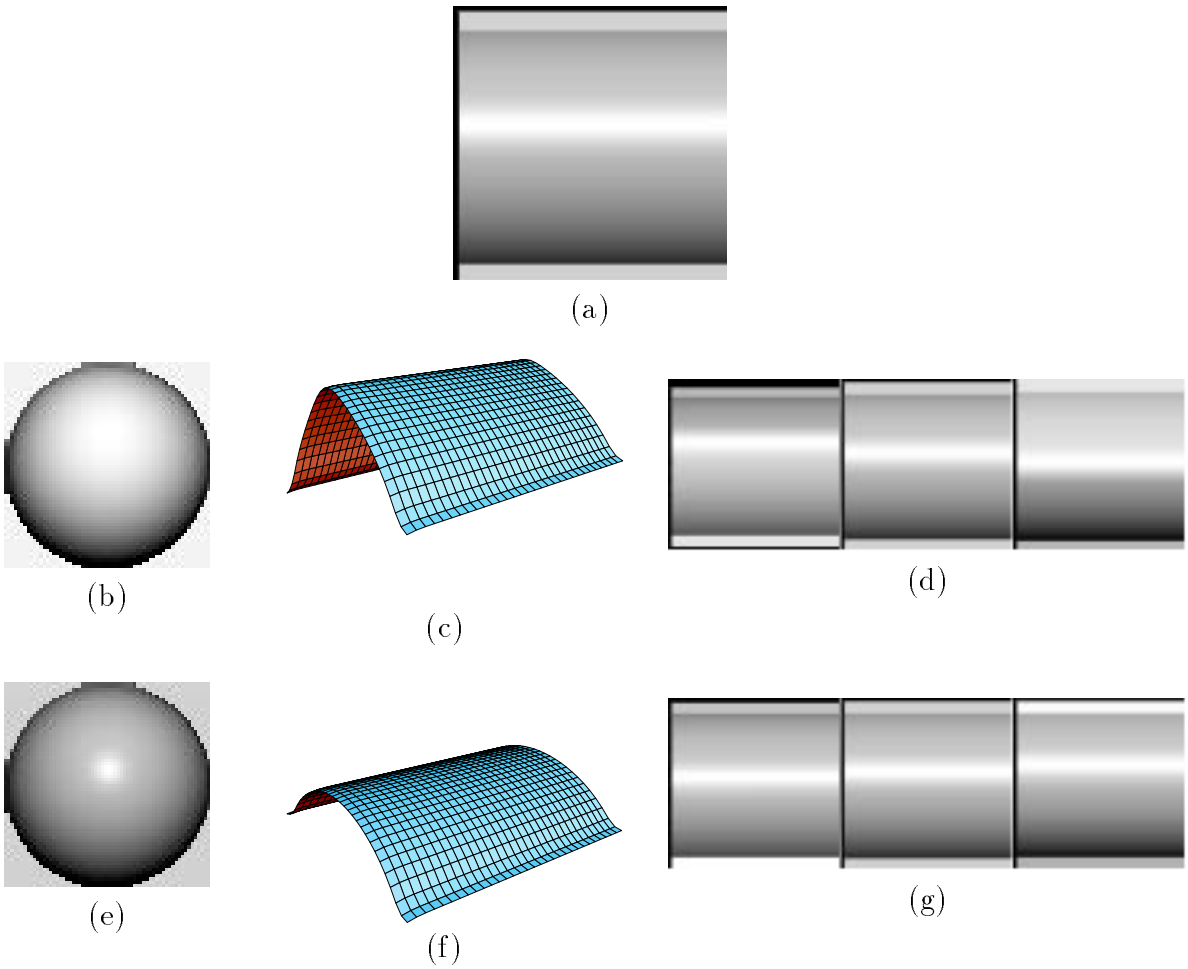


Figure 5: The image, (a), can be accounted for in two different ways. The shape (c) and the Lambertian reflectance function shown in (b) will create the image (a), as will the shape (f) and a shiny reflectance function (e). We can distinguish between the shiny and Lambertian explanations for (a) if we imagine rotating each shape. (d) and (g) show each shape at three different orientations. The image made from the shiny reflectance function, (e), changes only a little, while the parallax caused by the rotation of the tall shape of the Lambertian solution causes a larger image change. The reflectance function of (e) provides more angles over which the image looks nearly the same. If all viewpoints are equally likely, and the shapes and reflectances of (c) and (f) are equally likely to occur in the world, then (f) should be more likely than (c). The scene probability equation makes this precise.

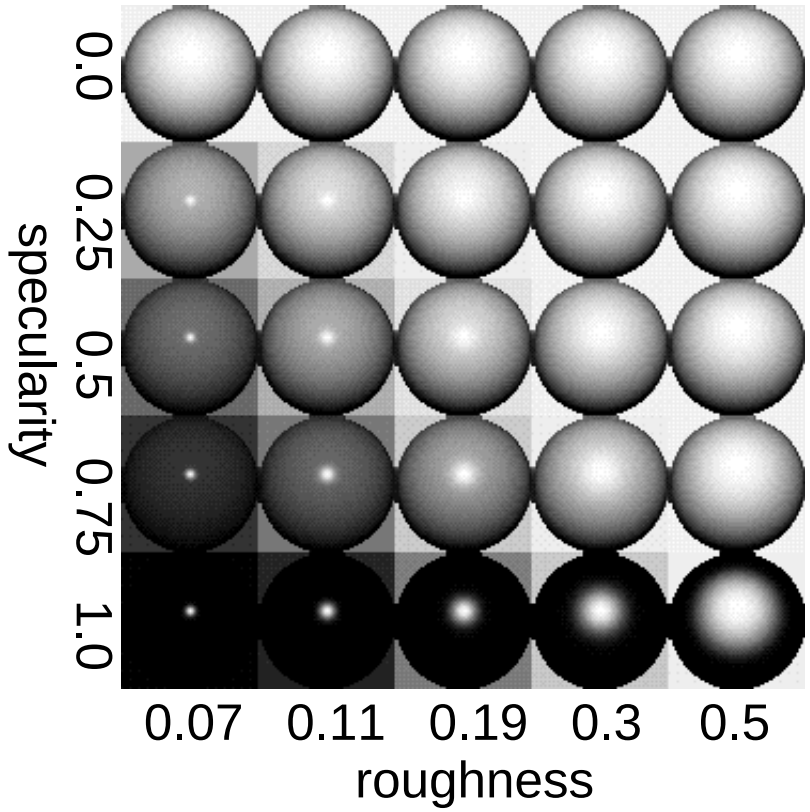


Figure 6: Key to reflectance function parameters of Fig. 7. Reflectance functions are displayed as they would appear on a hemisphere, lit in the same way as Fig. 7 (a) and (c). The ratio of diffuse to specular reflectance increases in the vertical direction. The surface roughness (which only affects the specular component) increases horizontally. The sampling increments are linear for specularity and logarithmic for roughness.

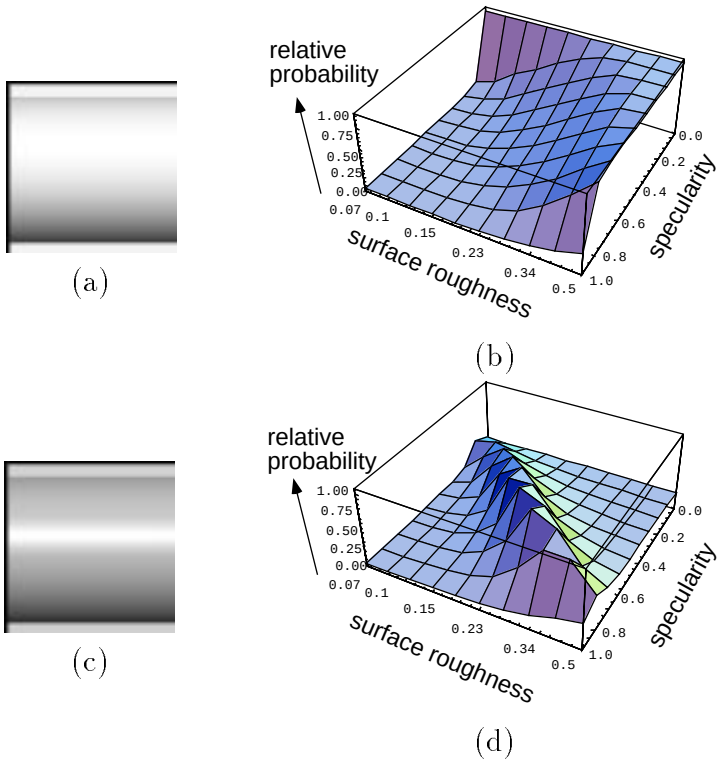


Figure 7: (a) Input image. (b) Probability that the observed image (a) was created by each reflectance function and corresponding shape. The probabilities are highest for the reflectance functions which look like the dull cylinder. See Fig. 6 for a visual guide to the reflectance function parameters of plots (b) and (d). (c) Input image. (d) Probability that (c) was created by each reflectance function and shape. The probabilities are highest for the reflectance functions which look like (see Fig. 6) the shiny cylinder. All reflectance functions were assumed to be equally likely and all can account for the image data equally well. The distinctions between reflectance functions came from the generic view term of the scene probability equation.

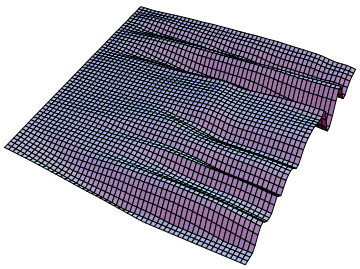
3.2 Why the Prior Probability Is Not Enough

Fig. 8 shows an example where both the fidelity and prior probability terms favor a perceptually implausible explanation. The generic view term biases the probabilities toward the plausible explanation. Fig. 8 (a) shows an image, and (b) and (c) are two possible explanations for it. (The vertical scale of (b) is exaggerated by 7). (We made this example by construction. Gaussian random noise at a 7 dB signal to noise ratio was added to (e) to make (a). (b) was found from (a) using a shape from shading algorithm, assuming constant surface height at the left picture edge [2]. We evaluated the probabilities of (b) and (c) assuming both generic object pose and generic lighting direction. The strength of a prior preference for smooth surfaces is arbitrary and none was included in the final densities. The actual noise variance was used for σ^2 in the fidelity term of Eq. (11), although a wide range of assumed variances would give the preferences described here).

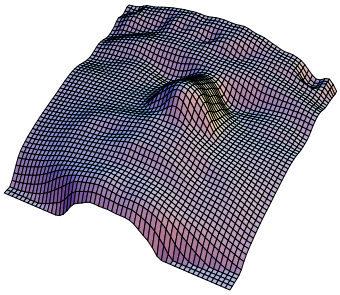
Perceptually, Fig. 8 (c) seems like a better explanation of (a), even though it doesn't account for all the noise. The fidelity term of the scene probability equation Eq. (11) favors Fig. 8 (b). Without the generic view term, the only term left to bias the probability of an interpretation is the prior probability. A typical prior is to favor smooth surfaces, which again would favor the peculiar shape (b), since (b) is much smoother than (c), as measured by the squared second derivatives of the surfaces.

What is missing from that framework is some way to penalize the precise alignment between the light source and the object which is required the get the image (d) from the shape (b). While the shape (c) doesn't account for the noise, it gives an image that is more stable with respect to object or lighting rotations. By the criterion of the scene probability equation, (c) has a higher overall probability density than interpretation (b). Fig. 8 (c) also corresponds more closely to one's visual perception of the object. Using the scene probability equation, we can recognize interpretations which are less faithful to the image data, yet more likely to have created the observed image.

(a)



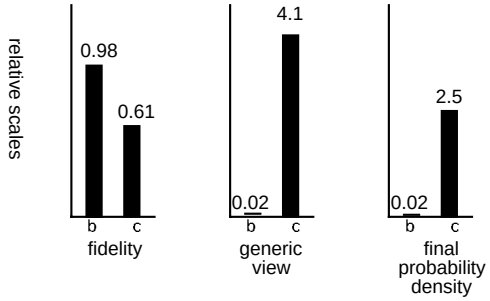
(b)



(c)

(d)

(e)


relative scales
0.98
0.61
b
c
fidelity
4.1
0.02
b
c
generic view
2.5
0.02
b
c
final probability density

(f)

Figure 8: The image in (a) can be explained by the surface (b), lit from the left (vertical scale exaggerated by 7). When rendered, that shaded shape accounts well (d) for the input image, but the object and the light source must be precisely positioned. Shape (c) is another possible interpretation. When lit from above, it does not account for the noise of the input image, as shown in (e). Both the fidelity and the typical prior of surface smoothness favor interpretation (b) over the shape (c). It is the generic view term of the scene probability equation which lets us penalize the precise alignment required for shape (b) to produce image (d). The relative values of the fidelity and generic view terms are shown in (f). The large generic view contribution of shape (c) gives it a higher overall probability, in accord with the human perception in preferring interpretation (e).

4 Relationship to Aspect Graphs

Aspect graphs [31, 32] also incorporate viewpoint information in a visual analysis. It is useful to compare the aspect graphs with our Bayesian approach. To make an aspect graph, one divides all possible views of an object into categories. A view is accidental if it lies "between" categories.

We re-draw the aspect graph Figure 9 (a) as a plot of image category as a function of viewpoint position in (b). This corresponds to a plot of the image data $f(x, \beta)$ as a function of the generic variable x in (c).

Knowing $P_{\vec{x}}(\vec{x})$, we can use Eq. (6) to find the probability of observing a particular aspect category, Fig. (9) (d). We integrate the distance in x over which the output $f(x)$ is that category. We do the same when we evaluate Eq. (6) analytically to arrive at the scene probability equation, Eq. (11). The analytic approximation preceeding Eq. (11) would fail for the categorical descriptions, for which the image category is piecewise constant over object pose. Then the denominator of the analytic approximation goes to zero. Thus we see that the aspect graph and scene probability equation approaches are complementary. They can perform a similar calculation of view probabilities, but in complementary domains.

5 Comments about the scene probability equation

We derived the scene probability equation for the case of Gaussian observation noise. There is some physical justification for this, since Gaussian noise is the limiting case of a sum of independent random variables. However, to study the effect of assumed noise distribution, we can consider a different noise distribution for the problem of Figure 3. We consider uniformly distributed zero mean observation noise of amplitude γ, identically distributed at each pixel. Then $P(\hat{\vec{y}} \mid \vec{\beta}, \vec{x})$ will be zero when any pixel of the rendered scene, $f(\vec{\beta}, \vec{x})$, differs from the observed image $\hat{\vec{y}}$ by more than γ. Otherwise $P(\vec{y} \mid \vec{\beta}, \vec{x}) = c$, a constant. Using the first order Taylor series term of Eq. (8), we find that the value of the integral of Eq. (6) for $P(\hat{\vec{\beta}} \mid \vec{y})$ varies inversely with the magnitude of largest derivative at any pixel value. We calculated this for the example of Figure 3; the results are shown in Figure 10.

As with guassian noise, the result favors the bump and dimple explanations. However, for this case of uniform noise, the different shape explanations are closer to equally probable. Thus we see that the assumed observation noise can change posterior probabilities. The uniform observation noise of Fig. 10, which implies an image metric based on the single pixel of largest difference, may be a poor choice.

The Gaussian noise which we assumed to derive the scene probability equation has drawbacks, as well. For independent, additive Gaussian noise, the probability that the difference between two images was caused by noise depends only on the sum of the squared differences between the images. Such a mean square error image distance metric is well known to be an imperfect measure of perceptual distance between images (e.g. [44]). The posterior probabilities of shapes 2–4 of Figure 3 (e) were larger than might be expected. Our Gaussian noise model may account for that. The long, linear structures of the difference images of Figure 3 (c) are extremely visible to people, yet not particularly improbable to the noise model. A more perceptually based image distance metric may assign greater penalty to shapes 2–4. For example, one might expect improved results by scaling the variance of the gaussian noise by a measure of the local image contrast. This would make image changes more visible in low-contrast regions than in high-contrast regions, analogously with human perception. This would give additional penalty to the tube shapes of Figure 3 (b) , since, when the light source moves, they introduce small changes in low-contrast regions of the image.

There is a potential computational concern regarding the scene probability equation. The approximation we made to Eq. (6) requires that we find $\vec{x}_0$, the value of $\vec{x}$ which minimizes $\|\vec{y} - \vec{f}(\vec{x}, \vec{\beta})\|^2$. This may be simple for some generic variables. It was for our examples of object pose or illumination direction, and it is for problems with bilinear structure [8]. There could be cases, however, where this was a non-trivial optimization problem.

Finally, there may be cases where the denominator of the generic view term of the scene probability equation goes to zero when image derivatives with respect to the generic variables go to zero. This indicates that the approximation for the integral in Eq. (6) no longer holds. The alternative formulation of the next section avoids this problem.

6 Relationship to Loss Functions

Another point of view from which to examine this work is that of Bayesian loss functions. I came to this viewpoint in the course of joint work with D. Brainard on a Bayesian approach to color constancy [8]. Independently, A. Yuille and H. Bulthoff [51] have applied a related loss function analysis.

In the scene probability equation, we split world parameters into generic and scene parameters. We indicate that we do not care about the values of the generic variables by integrating the joint posterior over those variables. However, sometimes we may be interested in estimating the generic variables to rough accuracy. Bayesian loss functions provide a framework in which to specify these desired accuracies precisely.

Suppose we have a posterior distribution $P(\vec{x}|\vec{y})$ on a vector variable $\vec{x}$ given the observation $\vec{y}$. In Bayesian decision theory (e.g., [1]), one defines a *loss function*, $L(\vec{x},\vec{a})$, which is the penalty for guessing $\vec{a}$ when the real value of the variable to be estimated is $\vec{x}$. From the posterior $P(\vec{x}|\vec{y})$, one can calculate the *posterior expected loss*, $\bar{L}(\vec{a},\vec{y})$, for making the decision $\vec{a}$, conditioned on the observation data $\vec{y}$, as

$$\bar{L}(\vec{a},\vec{y}) = \int_{\vec{x}} L(\vec{x},\vec{a})\, P(\vec{x}|\vec{y})\, d\vec{x}. \tag{15}$$

An observer will want to know what value of $\vec{a}$, call it $\hat{\vec{x}}$, minimizes the expected loss.

Figure 11 shows four loss functions of particular interest, illustrated for a hypothetical 2-dimensional parameter space. All the loss functions we consider are of the form $L(\vec{x},\vec{a}) = L(\vec{x}-\vec{a})$, and the expected loss Eq. (15) is a convolution of the loss function with the posterior.

A quadratic loss function, $L(\vec{x},\vec{a}) = |\vec{x}-\vec{a}|^2$, Fig. 11 (a), has many mathematically appealing properties. For example, the center of mass of the probability is the unique global minimum of the expected loss. However, its penalty increases without bound. It seems that a loss function for the task of perception ought to give equal penalty to all obviously wrong answers.

The commonly used MAP estimator corresponds to a minus delta function loss, $L(\vec{x},\vec{a}) = \delta(\vec{x}-\vec{a})$. Recall that MAP estimation uses the maximum of the posterior distribution as the estimate for $\vec{x}$. This loss function says every guess which is not correct to infinitely high precision carries the same

loss. Only *exactly* the right answer counts. Such a requirement for accuracy is not realistic for perception problems.

Figure 11 (c) shows a loss function related to the generic view approach of this paper. We let the parameter vector $\vec{x}$ be composed of a scene parameter vector, $\vec{x}_\beta$, stacked on top of a generic variable vector, $\vec{x}_x$ (these were $\vec{\beta}$ and $\vec{x}$, respectively, in the notation of the other sections). The scene probability equation implicitly assumes a constant loss over all values of the generic parameters. Thus $L(\vec{x}, \vec{a}) = F(\vec{x}_\beta - \vec{a}_\beta)$, where F is some function of the scene parameters. To see how this assumption was made, note that in Eq. (6) we integrated over all values of the generic variables. This corresponds to the integration of the expected loss integral of Eq. (15) where $L(\vec{x}, \vec{a})$ is constant over the generic parameters. For the generic view work, we did not specify the loss function $F(\vec{x}_\beta - \vec{a}_\beta)$ to be applied to the scene parameters. Figure 11 (c) shows a delta function loss in that dimension.

Our approximation of Eq. (8) means that we only evaluate the integral Eq. (15) over a local area of the generic variable parameter space. Thus in practise, we only consider the effect of a small variation in the generic variables.

A potentially appealing compromise between these loss functions is what we call a minus "local mass" loss function [8]. We define

$$L(\vec{x}, \vec{a}) = -e^{\frac{-1}{2\mu^2}(\vec{x}-\vec{a})^2}. \tag{16}$$

(Yuille and Bulthoff independently proposed the same loss function [51]). This function rewards getting approximately the right answer, and penalizes all wrong answers beyond several standard deviation with essentially equal penalty. The scalar variance μ^2 in Eq. (16) easily generalizes to a matrix.

The local mass function can offer advantages over the generic view loss function. It does not require the separation of variables into generic and scene parameters; both are explicitly estimated. We performed an asymptotic expansion to evaluate the integral of Eq. (6). A similar expansion may be used to find the expected loss for the local mass loss function. The asymptotic expansion used for Eq. (6) does not hold when the derivatives $\vec{f}'$ and $\vec{f}''$ go to zero, but the analogous approximation for the local mass loss function does not have this problem. Substituting Eqs. (5), (7), and (16) into Eq. (15), redefining $\vec{x}$ to subsume $\vec{\beta}$ and assuming σ small [27, 5], we have for the

expected loss:

$$\bar{L}(\vec{x},\vec{y}) = k\ \exp\left(\frac{-\|\vec{y}-\vec{f}(\vec{x}_0)\|^2}{2\sigma^2}\right)\ P_{\vec{x}}(\vec{x}_0)\ \frac{1}{\sqrt{\det(\boldsymbol{C})}} \tag{17}$$

where

$$C_{ij} = \vec{f'}_i \cdot \vec{f'}_j - (\vec{y}-\vec{f}(\vec{x}_0)) \cdot \vec{f''}_{ij} + \frac{\sigma^2}{\mu^2}\delta_{ij}, \tag{18}$$

and δ_{ij} are the elements of the identity matrix. Note that the above equations are nearly the same as the scene probability equation, Eq. (11), except that $\vec{x}$ now refers to a combined vector of generic and scene parameters. Also, here, $\vec{x}_0$, a function of $\vec{x}$, is

$$\vec{x}_0 = \arg\min_{\vec{\xi}}[\|\vec{y}-\vec{f}(\vec{\xi})\|^2 + \frac{\sigma^2}{\mu^2}\|\vec{\xi}-\vec{x}\|^2] \tag{19}$$

The constant $\frac{\sigma}{\mu}$ in Eq. (18) allows for use of Eq. (17) even in cases where the image derivatives $\vec{f'}$ and $\vec{f''}$ are zero. Eq. (17) allows explicit estimation of generic variables as well as scene parameters, since they are both combined into the variable $\vec{x}$. Figure 12 illustrates graphically the relationship between the integrations used to obtain scene probability equation, Eq. (11), and the expected local mass loss, Eq. (17). The two equations reflect the similarities shown graphically.

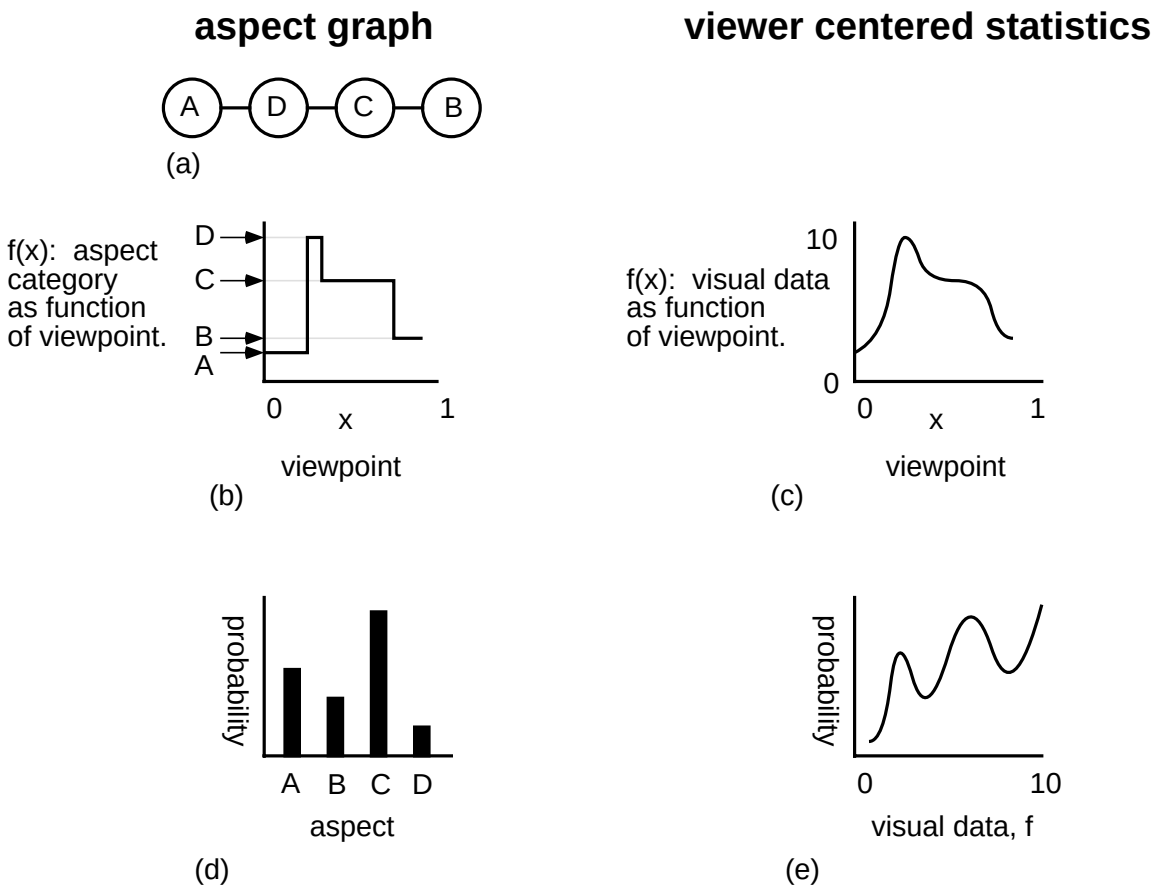


Figure 9: Comparison of aspect graphs and the present approach, for 1-d examples. (a) Hypothetical aspect graph of an object in a world where the viewpoint can only translate in one dimension. (b) The function relating viewpoint position, x, with aspect category is piecewise constant. In our continuous valued approach, the visual data is treated as a continuous function of viewpoint, x, plotted in (c). (d) Assuming generic viewpoint, the length of the viewpoint variable x for which $f(x)$ yields a given aspect category gives the probability of that aspect category. (e) For the scene probability equation, the local slope of $\frac{df(x)}{dx}$ influences the probability of the observation f.

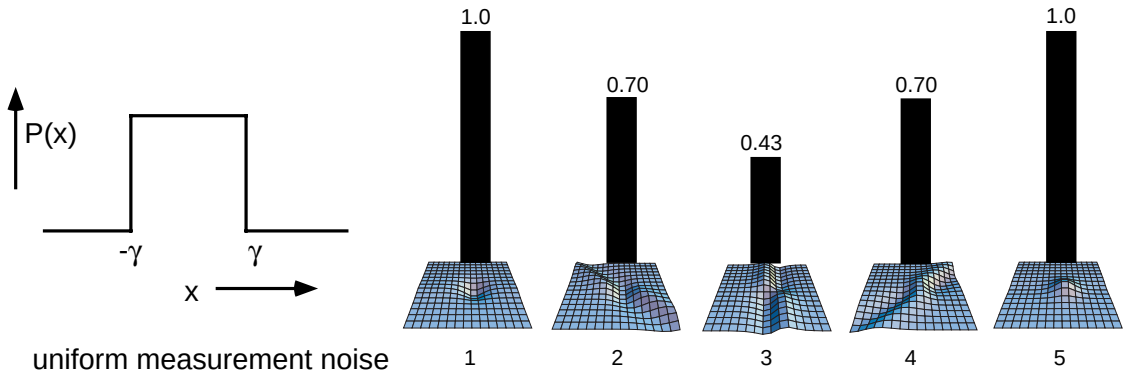


Figure 10: Relative probabilities of the 5 shapes of Figure 3, under the assumption of uniformly distributed observation noise. (In the analysis of Figure 3, we assumed Gaussian noise). For the model of uniform noise, the probability is based on the single pixel with the highest intensity derivative with respect to light direction. While favoring the bump and dimple interpretations, this noise model provides less discrimination between the shapes than does the gaussian noise model. Under the assumption of uniform noise, these probabilities are set by the single pixel with the largest intensity derivative with respect to the generic variable, in this case lighting direction. Discrimination between the shapes might be increased if noise detectability were made to be a function of local image contrast, as described in the text.

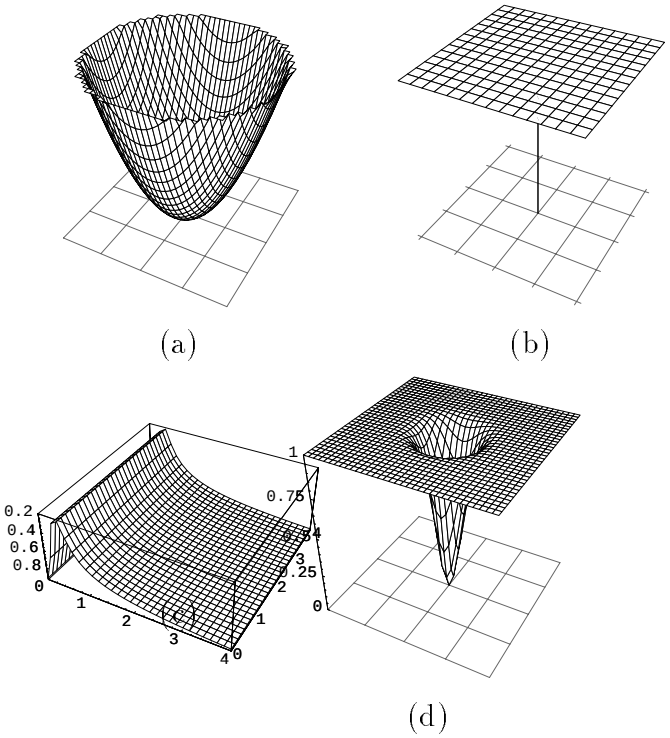


Figure 11: Four loss functions for a 2-d parameter space. (a) Quadratic loss function. The loss becomes unbounded as the error increases. (b) Minus delta function loss. A constant penalty is incurred for all but *exactly* the right answer. (c) Our generic view assumption was equivalent to a constant, extended loss function in the generic variable direction, such as shown here. In this chapter we did not specify the loss function in the scene parameter direction, which is shown as a negative δ function in the figure. (d) Minus "local mass" loss function. Small deviations from the correct answer are rewarded. Incorrect answers are penalized with a saturating penalty strength.

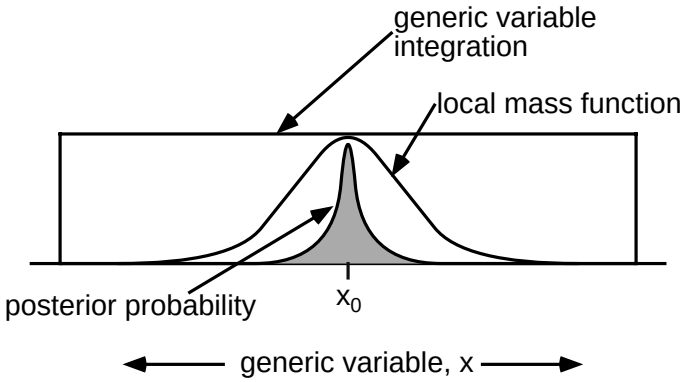


Figure 12: Similarity between marginalization of generic variable, used to derive Eq. (11), and the expected loss for a minus local mass loss function, used to derive Eq. (17). The shaded curve is the posterior probability. Typically, for a given scene interpretation, it will be high only over a narrow range of the generic variable, as shown. The marginalization integral over the generic variable measures the probability mass under the wide rectangle shown. The expectation integral for the local mass loss function measures the area under the gaussian function shown. For local mass functions much broader than the posterior probability spikes, these can give nearly the same answer.

7 Summary

It is often the case that there is more than one explanation for given visual data. In those cases, assumptions about the world must be used to break ties. The generic view assumption has commonly been used to categorize scene interpretations as either "generic" or "accidental". Here we apply this powerful assumption in a complementary domain, exploiting it to make quantitative judgements about scene parameters. We show how to invoke the assumptions of generic viewpoint, lighting, or object pose to estimate parameter values of scene interpretations. This approach removes some of the decision-making burden from the prior assumptions about the objects being estimated, for example, surface shape.

The input can be greyscale images or other visual data. We divide world parameters into two types, scene parameters, which we want to estimate, and generic parameters, which we do not. Following an approach taken in other application areas of Bayesian statistics, we write the image in a Taylor series expansion of the generic variables and integrate over the probability densities of the generic variables. The resulting *scene probability equation* gives the probability of a set of scene parameters, given an observed image. It has three major terms:

> a *fidelity term*, which requires that the scene parameters explain the observed visual data;
>
> the *prior probability*, which accounts for prior assumptions about the scene parameters;
>
> the *generic view term*, which quantifies how accidental our view of a particular scene is. It indicates the chance that a given scene would have presented us with the observed image.

We show various applications to shape from shading. The scene probability equation gives the probability of different reflectance functions and shapes for a given image. We assign relative probabilities to different shapes, each of which would generate the observed image, for different assumed light directions. The generic view term in the scene probability is important; one can have a shape from shading solution which is faithful to the data, but unlikely, and one which is less faithful but more likely.

The loss functions of Bayesian decision theory encompass our generic view framework as a special case. We present a modified version of the scene probability equation which applies for the case of a minus "local mass" loss function. This version of the scene probability equation avoids some problems where image derivatives go to zero and allows estimation of generic as well as scene parameters.

This Bayesian approach may have many applications in vision. The scene probability equation derived in this paper ranks the relative probability densities of different scene interpretations. From such an equation, one may derive algorithms which find an optimum scene interpretation. This may result in more powerful or more accurate algorithms for such problems as shape from shading, motion analysis, or stereo.

Acknowledgements

For helpful discussions and suggestions, thanks to: Ted Adelson, David Brainard, David Knill (especially about parameter estimation), David Mumford, Ken Nakayama, Sandy Pentland, Eero Simoncelli, Richard Szeliski, and Alan Yuille (who corrected an error). Three reviewers (Knill, Yuille, and a third) gave insightful reviews. Part of this research was performed at the MIT Media Laboratory and was supported by a contract with David Sarnoff Research Laboratories (subcontract to the National Information Display Laboratory) to E. Adelson.

References

[1] J. O. Berger. *Statistical decision theory and Bayesian analysis.* Springer-Verlag, 1985.

[2] M. Bichsel and A. P. Pentland. A simple algorithm for shape from shading. In *Proc. IEEE CVPR*, pages 459–465, Champaign, IL, 1992.

[3] I. Biederman. Human image understanding: recent research and a theory. *Comp. Vis., Graphics, Image Proc.*, 32:29–73, 1985.

[4] T. O. Binford. Inferring surfaces from images. *Artificial Intelligence*, 17:205–244, 1981.

[5] N. Bleistein and R. A. Handelsman. *Asymptotic expansions of integrals*. Dover, 1986.

[6] G. E. P. Box and G. C. Tiao. A Bayesian approach to the importance of assumptions applied to the comparison of variances. *Biometrika*, 51(1 and 2):153–167, 1970.

[7] G. E. P. Box and G. C. Tiao. *Bayesian inference in statistical analysis*. John Wiley and Sons, Inc., 1973.

[8] D. H. Brainard and W. T. Freeman. Bayesian method for recovering surface and illuminant properties from photosensor responses. In *Proceedings of SPIE*, volume 2179, San Jose, CA, February 1994. Human Vision, Visual Processing and Digital Display V.

[9] M. J. Brooks, W. Chojnacki, and R. Kozera. Impossible and ambiguous shading patterns. *Intl. J. Comp. Vis.*, 7(2):119–126, 1992.

[10] M. J. Brooks and B. K. P. Horn. Shape and source from shading. In B. K. P. Horn and M. J. Brooks, editors, *Shape from shading*, chapter 3. MIT Press, Cambridge, MA, 1989.

[11] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. In *SIGGRAPH-81*, 1981.

[12] S. J. Dickinson, A. P. Pentland, and A. Rosenfeld. 3-d shape recovery using distributed aspect matching. *IEEE Pat. Anal. Mach. Intell.*, 14(2):174–198, 1992.

[13] R. A. Fisher. *Statistical methods and scientific inference*. Hafner, 1959.

[14] W. T. Freeman. Exploiting the generic view assumption to estimate scene parameters. Vision and Modeling Technical Report 196, The Media Lab, MIT, 20 Ames St., Cambridge, MA 02139, 1992.

[15] W. T. Freeman. Exploiting the generic view assumption to estimate scene parameters. In *Proc. 4th Intl. Conf. Computer Vision*, pages 347 – 356, Berlin, Germany, 1993. IEEE.

[16] W. T. Freeman. The generic viewpoint assumption in a framework for visual perception. *Nature*, 368(6471):542–545, April 7, 1994.

[17] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Pat. Anal. Mach. Intell.*, 6:721–741, 1984.

[18] E. Grimson. Binocular shading and visual surface reconstruction. *Comp. Vis., Graphics, Image Proc.*, 28:19–43, 1984.

[19] S. F. Gull. Bayesian inductive inference and maximum entropy. In G. J. Erickson and C. R. Smith, editors, *Maximum Entropy and Bayesian Methods in Science and Engineering*, volume 1. Kluwer, 1988.

[20] S. F. Gull. Developments in maximum entropy data analysis. In J. Skilling, editor, *Maximum Entropy and Bayesian Methods, Cambridge*, pages 53–71. Kluwer, 1989.

[21] J. E. Hochberg. *Perception.* Prentice-Hall, Inc., Englewood Cliffs, N. J., 1978.

[22] B. K. P. Horn. Height and gradient from shading. Technical Report 1105, MIT Artificial Intelligence Lab, MIT, Cambridge, MA 02139, 1989.

[23] B. K. P. Horn and M. J. Brooks. *Shape from shading.* MIT Press, Cambridge, MA, 1989.

[24] B. K. P. Horn, R. Szeliski, and A. Yuille. Impossible shaded images. *IEEE Pat. Anal. Mach. Intell.*, 15(2):166–170, 1993.

[25] B. K. P. Horn, R. J. Woodham, and W. M. Silver. Determining shape and reflectance using multiple images. Technical Report 490, Artificial Intelligence Lab Memo, Massachusetts Institute of Technology, Cambridge, MA 02139, 1978.

[26] H. Jeffreys. *Theory of probability.* Oxford, Clarendon Press, 1961.

[27] H. Jeffreys. *Asymptotic approximations.* Oxford, Clarendon Press, 1962.

[28] A. Jepson and W. Richards. What makes a good feature? In L. Harris and M. Jenkin, editors, *Spatial Vision in Humans and Robots.* Cambridge Univ. Press., 1992. See also MIT AI Memo 1356 (1992).

[29] R. A. Johnson. Asymtotic expansions associated with posterior distributions. *The annals of mathematical statistics*, 41(3):851–864, 1970.

[30] D. Kersten. Transparancy and the cooperative computation of scene attributes. In M. S. Landy and J. A. Movshon, editors, *Computational Models of Visual Processing*, chapter 15. MIT Press, Cambridge, MA, 1991.

[31] J. J. Koenderink and A. J. van Doorn. The internal representation of solid shape with respect to vision. *Biol. Cybern.*, 32:211–216, 1979.

[32] D. Kriegman and J. Ponce. Computing exact aspect graphs of curved objects: Solids of revolution. *Intl. J. Comp. Vis.*, 5:119–135, 1990.

[33] P. S. Laplace. *Theorie analytique des probabilites*. Courcier, 1812.

[34] Y. G. Leclerc and A. F. Bobick. The direct computation of height from shading. In *Proc. IEEE CVPR*, pages 552–558, Maui, Hawaii, 1991.

[35] D. V. Lindley. *Bayesian statistics, a review*. Society for Industrial and Applied Mathematics (SIAM), 1972.

[36] D. G. Lowe and T. O. Binford. The recovery of three-dimensional structure from image curves. *IEEE Pat. Anal. Mach. Intell.*, 7(3):320–326, 1985.

[37] D. J. C. MacKay. Bayesian interpolation. *Neural Computation*, 4(3):415–447, 1992.

[38] J. Malik. Interpreting line drawings of curved objects. *Intl. J. Comp. Vis.*, 1:73–103, 1987.

[39] K. Nakayama and S. Shimojo. Experiencing and perceiving visual surfaces. *Science*, 257:1357–1363, 1992.

[40] A. Pentland. Photometric motion. In *Proceedings of 3rd International Conference on Computer Vision*, 1990.

[41] A. P. Pentland. Linear shape from shading. *Intl. J. Comp. Vis.*, 1(4):153–162, 1990.

[42] T. Poggio, V. Torre, and C. Koch. Computational vision and regularization theory. *Nature*, 317(26):314–139, 1985.

[43] W. A. Richards, J. J. Koenderink, and D. D. Hoffman. Inferring three-dimensional shapes from two-dimensional silhouettes. *J. Opt. Soc. Am. A*, 4(7):1168–1175, 1987.

[44] W. F. Schreiber. *Fundamentals of electronic imaging systems.* Springer-Verlag, 1986.

[45] J. Skilling. Classic maximum entropy. In J. Skilling, editor, *Maximum Entropy and Bayesian Methods, Cambridge*, pages 45–52. Kluwer, 1989.

[46] R. Szeliski. *Bayesian Modeling of Uncertainty in Low-level Vision.* Kluwer Academic Publishers, Boston, 1989.

[47] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-posed Problems.* Winston, Washington, DC, 1977.

[48] D. Weinshall, M. Werman, and N. Tishby. Stability and likelihood of views of three dimensional objects. In *Proceedings of the 3rd European Conference on Computer Vision*, Stockholm, Sweden, May 1994.

[49] A. P. Witkin. Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17:17–45, 1981.

[50] R. J. Woodham. Photometric method for determing surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[51] A. L. Yuille and H. H. Bulthoff. Bayesian decision theory and psychophysics. In D. Knill and W. Richards, editors, *Visual Perception: Computation and Psychophysics.* Cambridge University Press, 1994.

[52] Q. Zheng and R. Chellapa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Pat. Anal. Mach. Intell.*, 13(7):680–702, 1991.