White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/10766/

Article:
Stone, J.V., Kerrigan, I. and Porrill, J. (2009) Where is the light? Bayesian perceptual priors for lighting direction. Proceedings of the Royal Society B: Biological Sciences, 276 (1663). pp. 1797-1804. ISSN: 1471-2954

https://doi.org/10.1098/rspb.2008.1635

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Where is the light?
Bayesian perceptual priors for lighting direction

File: bayes_estprior_PRSJan2009_ForWhiteRose.tex
JV Stone[1], IS Kerrigan[2], and J Porrill[1]
[1]Department of Psychology, University of Sheffield, England, S10 2TP.
[2]School of Psychology, University of Southampton.
Email: *j.v.stone@shef.ac.uk*
Full reference: Proceedings Royal Society (B), 276(1663) pp1797-804, 2009.

May 18, 2010

Abstract

Perception of shaded three-dimensional figures is inherently ambiguous, but this ambiguity can be resolved if the brain assumes that figures are lit from a specific direction. Under the Bayesian framework, the visual system assigns a weighting to each possible direction, and these weightings define a prior probability distribution for light source direction. Here, we describe a non-parametric maximum likelihood estimation method for finding the prior distribution for lighting direction. Our results suggest that each observer has a distinct prior distribution, with non-zero values in all directions, but with a peak which indicates observers are biased to expect light to come from above-left. The implications of these results for estimating general perceptual priors are discussed.

Introduction

Perception consists of interpreting two-dimensional (2D) retinal images of a three-dimensional (3D) world. The process of projecting a 3D scene onto a 2D retina necessarily discards information about the 3D structure of that scene. This makes it impossible, in principle, to deduce all of the 3D structure of a scene, and perception is therefore a classic example of an *ill-posed problem* (Poggio et al. 1985). However, even though such problems cannot be solved by deduction, acceptable solutions can be found using statistical inference. This involves using additional information, usually based on prior experience, to interpret 2D retinal images, where this additional information takes the form of heuristics (rules of thumb) or constraints (rules which exclude certain 'illegal' solutions).

Within the Bayesian framework, this extra information is realized in the form of prior distributions. For example, the image marked with a cross in Figure 1 can be interpreted either as convex or as concave. The particular perception evoked by this image depends only on the direction in which the light source is assumed to originate (Rittenhouse 1786, Brewster 1847, Oppel 1856, Kleffner & Ramachandran 1992). If the light source is assumed to originate from below then the image is interpreted as convex, but if the light source is assumed to originate from above then the image is interpreted as concave. As this is the usual interpretation made by human observers, it implies that we implicitly assume light originates from above. However, such demonstrations provide only a qualitative impression of where we assume the light source to be.

In reality, it is unlikely that human observers make the simplistic assumption that light comes only from above or below. More realistically, each observer assigns a probability to each possible light-source direction which may be based on prior experience of the directions in which light sources originate.

These probability values collectively define a prior probability density function, which can be visualized using a polar plot, where the radial distance in a given direction indicates the relative probability that the light originates from that direction (as in Figure 2c). In this paper, we show how it is possible to estimate the overall form of this prior, which, for reasons that will become obvious, we call the light-from-above prior. For the sake of clarity, note that we do not seek the prior for lighting direction, which could be obtained empirically, but the prior as used by a given observer.

Our general strategy is closely related to that described in Paninski (2006). However in the simulated experiment described by Paninski, the observer estimates a continuous parameter, and so each trial provides an equality constraint on the prior. Here we concentrate on the more common case in which the observer makes a forced choice, so that each trial provides a weaker, inequality constraint on the prior.

Results

The shape information in our images is a function of two parameters, the direction θ of the light source, and the 3D shape c of the imaged surface, which specifies whether the stimulus is concave $c = c_1$ or convex $c = c_0$. On each trial, the observer is presented with an image $\mathbf{x}$, and makes a binary response $r = 1$ if the stimulus appears concave, or $r = 0$ if the stimulus appears convex (see Appendix 1).

We assume that the observer's perceieved shape $\hat{c}$ of a shape c depends on two quantities: the posterior probability density function, and the loss function. First, the probability (density) that the shape has value c and that the light source is in direction θ given an image $\mathbf{x}$ defines the joint posterior probability density

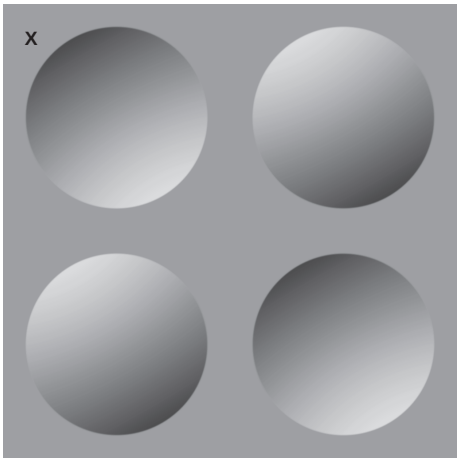


Figure 1: **Typical stimulus presented to observer on a single trial**. The observer's task is to indicate whether or not the quadrant marked with a cross (x) appears convex or concave. This response implicitly defines the perceived direction of the light source. For example, if the marked quadrant is perceived as convex then this implies that the light originates from the lower right (i.e. at about 300°), but if it is perceived as concave then this implies that the light originates from the upper left (i.e. about 120°).

function $p(c, \theta|\mathbf{x})$. Second, the 'cost' of perceiving a shape as $\hat{c}$, when it is actually c, is defined by the loss function $D(\hat{c}, c)$.

The observer's perception is assumed to correspond to the shape $\hat{c}$ which minimises the expected loss, where this expectation is taken over all possible values of θ and c

$$E = \int_c \int_\theta p(c, \theta|\mathbf{x}) D(\hat{c}, c)\, d\theta\, dc \quad (1)$$

$$= \int_c D(\hat{c}, c) \left[\int_\theta p(c, \theta|\mathbf{x})\, d\theta \right] dc. \quad (2)$$

Using Bayes' rule, the posterior is given by

$$p(c, \theta|\mathbf{x}) = p(\mathbf{x}|c, \theta) p(c, \theta)/p(\mathbf{x}), \quad (3)$$

where the observer's prior expectations about shapes and lighting directions define the joint prior distribution $p(c, \theta)$, and where the probability of the observed image for a given 3D shape and lighting direction defines the likelihood function $p(\mathbf{x}|c, \theta)$. The integral in square brackets in Equation (2) can now be re-written as

$$p(c|\mathbf{x}) = \frac{1}{p(\mathbf{x})} \int_\theta p(\mathbf{x}|c, \theta) p(c, \theta)\, d\theta, \quad (4)$$

so that the expected loss is

$$E = \int_c p(c|\mathbf{x}) D(\hat{c}, c)\, dc. \quad (5)$$

In fact, each stimulus $\mathbf{x}$ is consistent with only two lighting directions, $\theta_\mathbf{x}$ and $\overline{\theta}_\mathbf{x} = \theta_\mathbf{x} + 180°$. This implies that the likelihood $p(\mathbf{x}|c, \theta)$ is a delta function, which is zero except at $\theta = \theta_\mathbf{x}$ and $\theta = \overline{\theta}_\mathbf{x}$:

$$p(\mathbf{x}|c_0, \theta) = \delta(\theta - \theta_\mathbf{x}) \quad (6)$$

$$p(\mathbf{x}|c_1, \theta) = \delta(\theta - \overline{\theta}_\mathbf{x}). \quad (7)$$

Substituting Equation (6) in Equation (4) for $c = c_0$ yields

$$p(c_0|\mathbf{x}) = \frac{1}{p(\mathbf{x})} \int \delta(\theta - \theta_\mathbf{x}) p(c_0, \theta)\, d\theta, \quad (8)$$

$$= p(c_0, \theta_\mathbf{x})/p(\mathbf{x}). \quad (9)$$

If the observer assumes that the stimulus shape and the lighting direction are independent then the joint prior distribution $p(c_0, \theta_\mathbf{x})$ factorises to yield

$$p(c_0|\mathbf{x}) = p(c_0) p_\theta(\theta_\mathbf{x})/p(\mathbf{x}), \quad (10)$$

where $p_\theta(\theta_\mathbf{x})$ is the prior over lighting direction, and $p(c_0)$ is the prior for the shape c_0. A similar calculation for $c = c_1$ yields

$$p(c_1|\mathbf{x}) = p(c_1) p_\theta(\overline{\theta}_\mathbf{x})/p(\mathbf{x}). \quad (11)$$

Regardless of the value of $\theta_{\mathbf{x}}$ and $\overline{\theta}_{\mathbf{x}}$, each observer perceives the stimulus as either convex c_0 or concave c_1, and responds accordingly. Thus, together, $p(c_1)$ and $p(c_0)$ is a pair of co-determined observer-specific scalar priors, such that $p(c_1) + p(c_0) = 1$. We call the prior $p(c_1)$ the concavity preference for a given observer, which can be estimated using the same method (described below) for estimating the prior $p_\theta(\theta)$.

We choose the zero/one loss function to model the forced choice task, that is $D(\hat{c}, c) = 0$ for a correct decision, and $D(\hat{c}, c) = 1$ for an incorrect decision (Bishop 1996), the optimal decision rule under this loss function minimises the number of mis-classified stimuli. Substituting this loss function into Equation (5) we find that the observer should respond $r = 0$ (convex) if the log posterior ratio

$$\begin{aligned} L &= \log \frac{p(c_0|\mathbf{x})}{p(c_1|\mathbf{x})} && (12)\\ &= \log \frac{p(c_0)p(\theta_{\mathbf{x}})}{p(c_1)p(\overline{\theta}_{\mathbf{x}})} && (13)\\ &\geq 0. && (14) \end{aligned}$$

and the response should be $r = 1$ (concave) otherwise.

This deterministic rule would lead to the same decision for all presentations of a given stimulus. In order to model the stochastic character of human decision making, we follow a general suggestion of (Paninski 2006), and assume that our rule is stochastic (see Discussion). Specifically, we assume that the process (e.g. the observer's criterion) which compares the log-posterior probability $\log p(c_0|\mathbf{x})$ with $\log p(c_1|\mathbf{x})$ is subject to noise. In order to be clear about the implications of this, we define

$$\begin{aligned} L_0 &= 10 \log_{10} p(c_0)p_\theta(\theta_{\mathbf{x}}) && (15)\\ L_1 &= 10 \log_{10} p(c_1)p_\theta(\overline{\theta}_{\mathbf{x}}), && (16) \end{aligned}$$

and re-write Equation (14) as $L = L_0 - L_1$. We assume that the distribution of L_0 values is Gaussian with mean $\overline{L}_0$ and standard deviation σ, and that the distribution of L_1 values is Gaussian with mean $\overline{L}_1$, also with standard deviation σ. As L_0 and L_1 are both Gaussian with variance σ^2, L is also Gaussian with mean $\overline{L} = \overline{L}_0 - \overline{L}_1$ and variance $\sigma_L^2 = 2\sigma^2$. For simplicity, we assume that σ is the same for all lighting directions.

Note that we have chosen to measure the relative log-likelihood (sometimes called evidence) in decibels (dB) as suggested by (Jaynes 2003). This allows easy comparison of levels of evidence. For example, evidence of 3dB for an hypothesis means that it is about twice as likely than its alternative, and 10dB means that it is about ten times as likely. Jaynes has suggested that an evidence threshold of about 1dB is characteristic of many human judgments (Jaynes 2003).

We assume that the probability $P(c_0|\mathbf{x})$ of the observer perceiving a shape c_0 is described by the cumulative density function of a Gaussian with zero mean and variance σ_L^2

$$\begin{aligned} P(c_0|\mathbf{x}) &= \frac{1}{\sigma_L\sqrt{2\pi}} \int_{-\infty}^{L} e^{-\eta^2/(2\sigma_L^2)}\, d\eta && (17)\\ &= (1 + \text{erf}(L\sqrt{2}/\sigma_L))/2 && (18)\\ &= q, && (19) \end{aligned}$$

where q is defined for brevity.

For a given value of q, if the same stimulus is presented on n trials and if responses are independent across trials then the probability that the observer responds $r = 1$ (concave) on m of those n trials is

$$p(m|q) = C_{n,m}\, q^m (1-q)^{n-m}, \tag{20}$$

where $C_{n,m}$ is a binomial coefficient. For a given light direction, $C_{n,m}$ is constant, and so it does not affect the value $\hat{q}_i$ that maximises $p(m|q)$, and is omitted below.

We discretize the lighting direction into N values: $\theta_i : i = 1, \ldots, N$. For a given value of θ_i we present the stimulus n_i times, and record the number m_i of 'concave' responses, so that

$$p(m_i|q_i) = q_i^{m_i}(1-q_i)^{n_i - m_i}. \tag{21}$$

Thus, the n_i binary responses of a single observer to repeated presentations of the same stimulus are maximally consistent with the value $\hat{q}_i$ of q_i, which is the probability that the observer perceives the shape as concave when the lighting direction is θ_i.

When considered over all N lighting directions, and assuming independent noise, the probability of the vector $\mathbf{m} = (m_1, \ldots, m_N)$ for a given vector $\mathbf{q} = (q_1, \ldots, q_N)$ is

$$p(\mathbf{m}|\mathbf{q}) = \prod_i q_i^{m_i}(1-q_i)^{n_i - m_i}, \tag{22}$$

which is the the likelihood function of $\mathbf{q}$. The vector $\hat{\mathbf{q}}$ of $\mathbf{q}$ that maximises $p(\mathbf{m}|\mathbf{q})$ is the maximum likelihood estimate of the true value $\mathbf{q}^*$. Taking logs and multiplying by minus one transforms Equation (22) into the negative log likelihood function of $\mathbf{q}$

$$E_f = -\sum_i^N m_i \log q_i + (n_i - m_i)\log(1-q_i). \tag{23}$$

As both the prior distribution and the concavity preference are implicit in $\hat{\mathbf{q}}$, this provides an estimate $\hat{p}_\theta(\theta)$ of the true prior distribution $p^*_\theta(\theta)$, and an estimate $\hat{p}(c_1)$ of the true concavity preference $p^*(c_1)$.

As discussed later, the unknown value of the discrimination parameter σ_L means that, in practice, the prior is not completely determined by Equation (14) (see Discussion); but for the sake of brevity, we will refer to this as 'estimating the prior'.

Smoothing the prior: Unless the data set is very large, the prior distribution estimated by direct minimisation of E_f will not be very smooth. Smoothness of the prior probability for lighting direction is an important physical constraint which we can model by regularising the solution

$$E = E_f + \lambda^2 E_s, \tag{24}$$

where E_s is a measure of the smoothness of $p_\theta(\theta)$, and λ is proportional to the square of the expected angular scale over which the prior for lighting direction is expected to change. This regularization procedure can be thought of as specifying a 'prior for priors' (Paninski 2006).

Paninski suggests using the usual L_2 norm on the derivative of the prior to measure smoothness. A related measure which is more appropriate to this probabilistic situation (see Discussion) is the Fisher information, which measures the extent to which the prior $p_\theta(\theta)$ is localised, and which is a weighted version of the usual L_2 norm:

$$E_s = \mathrm{E}\left[\left(\frac{d \log p_\theta(\theta)}{d\theta}\right)^2\right] \quad (25)$$

$$= \sum_i \frac{(p_\theta(\theta_{i+1}) - p_\theta(\theta_i))^2}{p_\theta(\theta_i)\,\Delta\theta}. \quad (26)$$

In summary, for given values of the smoothing parameter λ, the values of the N elements of the discretized prior $p_\theta(\theta)$ and the concavity preference $p(c_1)$ can be estimated simultaneously as those values which minimise E (Equation 24). The value of λ was estimated using cross-validation (Bishop 1996) (see Appendix 2), and the MatLab minimisation procedure 'fminsearch' was used to find an estimates of $p^*_\theta(\theta)$ and $p^*(c_1)$.

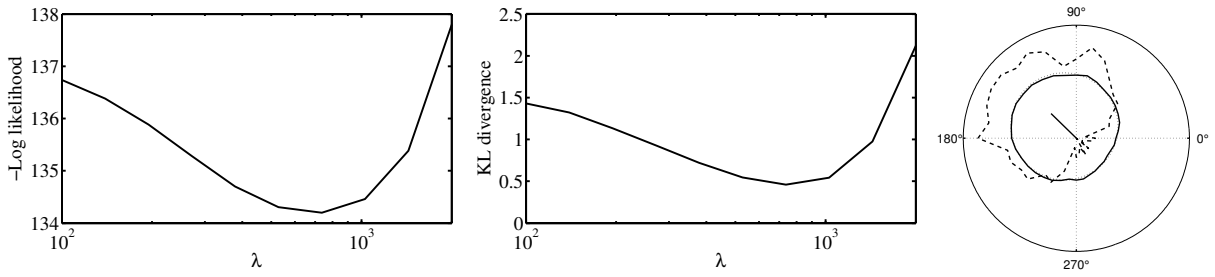
Results for Simulated Observer: In order to test our methods, we first analyzed data from a simulated observer with a known prior $p^*_\theta(\theta)$. The prior was defined as a von-Mises distribution (Fisher 1995) $p_\theta(\theta) = \exp(\kappa \cos((\theta - \mu))$ with location parameter $\mu = -45^\circ$ and dispersion parameter $\kappa = 0.33$. The value of the smoothing parameter λ has no explicit representation when generating data for the simulated observer, and cross-validation (Appendix 2) was used to find an estimate of $\hat{\lambda} = 700$, see Figure 2. This was then used with the known value of $\sigma = 1$ to estimate the simulated observer's prior $\hat{p}_\theta(\theta)$ for lighting direction and its concavity preference $\hat{p}(c_1)$. The concavity preference of this simulated observer had been defined as $p(c_1) = 0.5$, and was subsequently estimated as $\hat{p}(c_1) = 0.507$. The method also recovered an accurate estimate of the prior, as shown in Figure 2c.

Results for Human Observers: Using cross-validation (Appendix 2), the estimated value of the smoothing parameter was $\hat{\lambda} = 400$, see Figure 3. This was then used with $\sigma = 2$ to estimate each observer's prior $p_\theta(\theta)$, see Figure 4. In each case, the estimated prior is biased toward the upper left, in agreement with previous findings on group average data (Mamassian & Landy 2001). Thus the left-biases observed in each posterior in Figure 4 and in (Mamassian & Goutcher 2001), as well as the left and right biases reported in (Sun & Perona 1998, Adams et al. 2004), are probably due to a bias in each observer's prior, rather than a bias in the likelihod function. The estimated prior concavity preferences for all observers were within the range $\hat{p}(c_1) = 0.49 - 0.51$, compared with findings for the posterior in (Adams et al. 2004) (0.44) which used similar stimuli. Details of the experimental procedure are given in Appendix 1.

Discussion

When an observer is asked to report the concavity/convexity of a shape for a range of different lighting directions, the resultant set of responses (usually depicted as a polar plot) represents a sample from their posterior probability density function for shape. It is this sample from the observer's posterior which has been used in all previous experiments to provide estimates of observers' posterior for lighting direction.

The main contribution of this paper is a method for using this sampled posterior, in combination with a likelihood function and a loss function, to estimate the prior probability density function for lighting direction and the prior for concavity preference in individual observers. In order to achieve this, we assume plausible forms for the likelihood and loss functions. For the loss function, we assume that each observer attempts to minimise the number of mis-classified stimuli, an objective which corresponds to making responses consistent with the mode of the posterior probability density function. With regard to the likeli-



(a) (b) (c)

Figure 2: **Evaluating cross-validation method for estimating the value of the smoothing parameter** λ. Results for the simulated observer's data shown in Figure 2c for $\sigma = 1$. See Appendix 2.
(**a**). At each sampled value of λ, three quarters of the data were used to estimate the prior. Using this prior, the likelihood of the remaining quarter was evaluated using Equation (23), where $\sigma = 1$. This was repeated once for each of four disjoint quarters, and the mean of the four resultant likelihood functions is plotted here. The minimum with respect to the negative log likelihood corresponds to $\lambda \approx 700$.
(**b**). To evaluate the success of cross-validation for each sampled value λ_j of λ, the Kullback-Leibler (KL) divergence between the known prior $p^*_\theta(\theta)$ of this simulated observer and the prior $\hat{p}_\theta(\theta)$ obtained with λ_j and was calculated as $E_{KL} = 1000 \times \Delta\theta \sum \hat{p}_\theta(\theta_i) \log \hat{p}_\theta(\theta_i)/p^*_\theta(\theta_i)$, where $i = (0, 10, \ldots, 350)$, $\theta_i = 10 \times i$ and $\Delta\theta = 10^\circ$. The minimum with respect to the KL distance also corresponds to $\lambda \approx 700$, confirming that cross-validation chooses a value of λ which provides a good estimate of the true prior.
(**c**) **Estimating the light-from-above prior** $p_\theta(\theta)$ **for a simulated observer**. The lighting direction varies around the circle, and the probability that the stimulus was judged to be convex varies with distance from the origin. The graph shows 1) a sample from the posterior $p(c_0|\mathbf{x})$ as the proportion of convex responses (dashed), 2) the known prior (dotted), 3) the estimated prior (solid), based on the proportion of convex responses, as it would be for a human observer, 4) the mean vector (solid line), which is the mean of the prior (see Appendix 3). The direction of this vector indicates the bias in the prior, and its length shows the amount of bias (see Methods). The simulated observer was exposed to the same 36 lighting directions and the same number of trials per lighting direction (32) as the human observers used in the experiment described in the text, a discrimination parameter which was set at $\sigma^* = 1$ dB, and a concavity preference set at $p(c_1) = 0.5$. The value of the smoothing parameter estimated from cross-validation is $\lambda = 700$ (see a and b). Using $\lambda = 700$ and $\sigma = 1$, the concavity preference was estimated as $\hat{p}(c_1) = 0.507$.

hood function, each convexity/concavity response is consistent with one of two possible lighting directions, which effectively implies that the likelihood function is a delta function with non-zero values corresponding to these two lighting directions. This provides a posterior which is proportional to the prior for exactly two lighting directions and two shapes (convex/concave). An estimate of each observer's prior and concavity preference was then obtained by minimising a regularized (smoothed) version of the negative log likelihood of the sampled posterior.

Related Work. Research on motion perception explained the change in perceived speed that occurs at different levels of contrast by assuming a specific (Gaussian) form the speed prior (Weiss et al. 2002). Other researchers assume that the mean of the posterior coincides with the true stimulus value in a sensorimotor task (Körding & Wolpert 2004) or that, i) the log of the prior is a straight line, ii) the likelihood is Gaussian, and iii) mean of the posterior is the true mean (Stocker & Simoncelli 2006a). We make none of these

assumptions.

A parametric estimate of the lighting prior has previously been obtained (Mamassian & Landy 2001) under the assumption that it can be described by a 2-parameter von Mises distribution (see below).

The method described here is inspired by Paninski (2006). However, our method is different from Paninski's in a two key respects. First, the stochastic choice model assumes that the log posterior probabilities (and not posterior probabilities) are subject to additive Gaussian noise (Equation (17)). This has a number of advantages. i) The chosen value of σ corresponds naturally to a threshold value for the evidence (in the sense of log-posterior ratio) needed to obtain a given choice rate in the presence of encoding noise. ii) There are no problems of positivity in adding an unbounded noise contribution to probability values which should be positive. iii) The neural encoding of log probabilities has been shown to have a direct neural interpretation as an approximation to Poisson noise in neural populations (Gold & Shadlen 2001).

Second, we have replaced the L_2 regulariser used in (Paninski 2006) with Fisher information. This is more closely related to the probabilistic nature of the problem. Essentially, regularisation using Fisher information (Equation (26)) tries to satisfy the experimental constraints using the least localized prior density. By up-weighting the contribution for low probabilities (i.e. by $1/p_\theta(\theta)$) the Fisher regulariser takes account of the fact that small ripples in low probability regions are just as significant as larger ripples at higher probability values when the task requires a likelihood ratio judgment. There is inevitably a trade-off between form of the estimated prior and the nature of the smoothing function. However, because the Fisher regulariser seeks that prior with the least localized density, it can be interpreted as the regulariser of least commitment.

The Experimental Task. The design of the concavity-convexity task was chosen for a number of reasons. First, we have chosen a forced choice task. Experiments in which the observer provides an explicit estimate of lighting direction on each trial could provide more powerful constraints on the prior. However, asking observers to estimate lighting direction is an unnatural task, and is therefore likely to yield data that is both biased and noisy. Although we have seen that the forced choice experiment leaves some aspects of the prior unconstrained, it requires far fewer modelling assumptions than parameter estimation alternatives, and so the information that is obtained is more reliable.

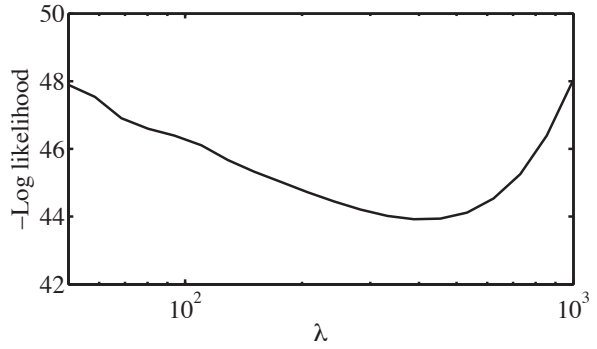


Figure 3: **Cross-validation**. Result for estimating the value of the smoothing parameter λ for observer *a* in Figure 4, with $\sigma = 2$. The minimum with respect to the negative log likelihood corresponds to $\lambda \approx 400$. This curve is typical of that obtained for other observers, and a value of $\hat{\lambda} = 400$ was therefore used for all human observers.

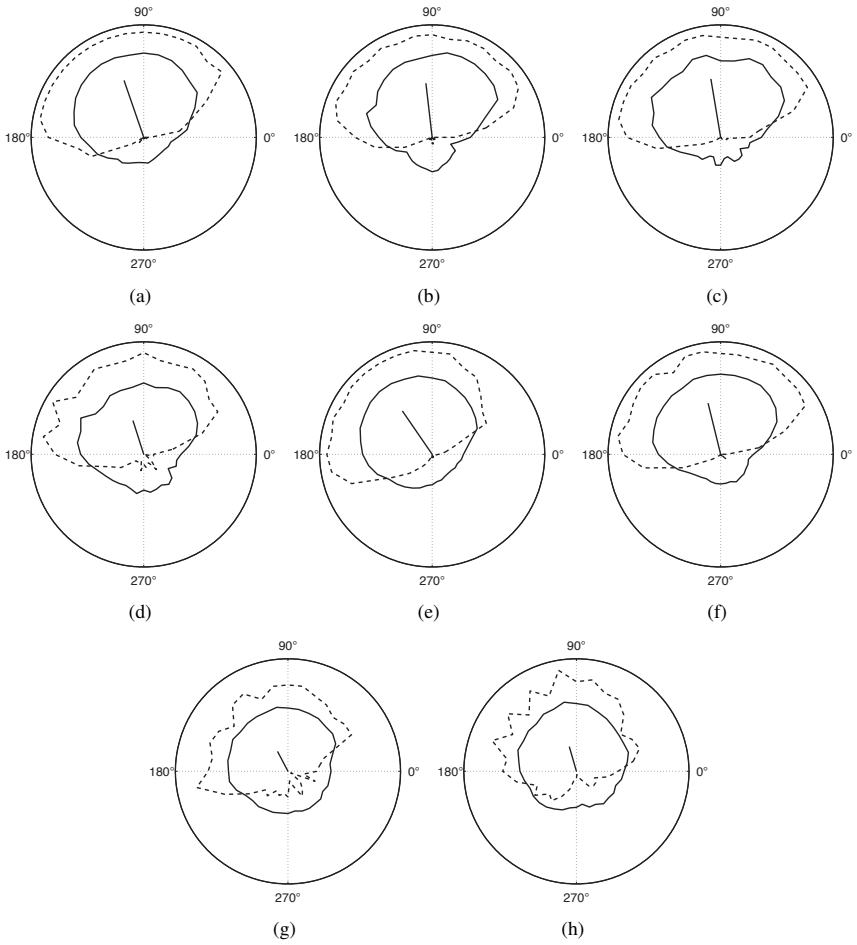


Figure 4: **Polar plots of estimated priors for 8 observers.**
Each graph shows the frequency of convex responses (dashed) as a function of light source direction. This is essentially a sample from the observer's posterior, and is used to estimate the prior (solid). For display purposes, the lengths of all mean vectors (solid line) have been scaled by the same factor across all graphs, and all graphs are drawn to the same scale (see Appendix 3). Note that all biases are to the left, with values of $20°, 7°, 9°, 18°, 34°, 14°, 28°$, and $16°$, respectively (mean $18°$). These results were obtained using all the data for each observer with $\lambda = 400$ and $\sigma = 2$.

Second, although the question of the modification of the prior by feedback is of great interest (Adams et al. 2004), no feedback was given here, and there is no correct response for the chosen stimuli. This is important because, in most applications, even a small number of trials with feedback reduces the dependence of the posterior on the prior to insignificance (Mele & Rawling 2004) (indeed, this 'washing out' property is often invoked to protect Bayesian methods from the consequences of choosing incorrect priors). In our experiment, neither the posterior nor the prior can be updated as a consequence of feedback. It is generally assumed that exposure to a biased population of stimuli (e.g. exposure to mainly concave stimuli) induces a shift in the prior. However, this appears to be the case only if feedback is given to correct the interpretation

of ambiguous stimuli. Observers adapt their visual interpretation of stimuli like those in Figure 1 provided they are given haptic feedback of those stimuli (Adams et al. 2004). Moreover, this adaptation was found to affect performance on a different (lightness judgment) task, which required an assumption regarding light direction, indicating a shift in the mean of the light-from-above prior. From a statistical perspective, this makes sense. Decisions based on a series of measurements with corrective feedback are initially based mainly on prior expectations. However, the corrective feedback can be used to update the prior, making future decisions more reliable, as in the classic Kalman filter (Kalman 1960). However, exposure to a biased population of stimuli without feedback induces after-effects in the opposite direction to that predicted by a shift in the prior. These after-effects are consistent with a change in the likelihood function and not in the prior (Stocker & Simoncelli 2006b). In our experiment, observers were exposed to an unbiased population of stimuli and received no feedback. Given the above considerations, this suggests that the prior and likelihood were not affected by the stimuli, and were reasonably constant throughout the experiment.

Third, we have used stimuli which are essentially noise free. Many visual tasks have unavoidable sensory noise, and when this is not the case experimenters have added artificial noise, specifically in order to allow a Bayesian analysis. By virtually eliminating this sensory noise in a very simple task, we have ensured that

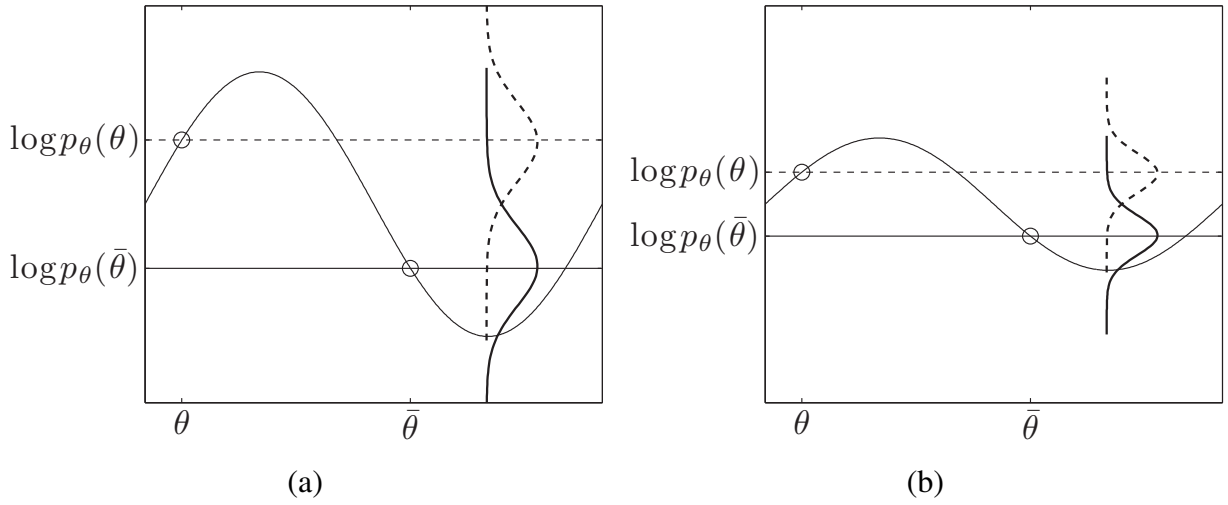


Figure 5: **The discrimination parameter σ is undetermined.**
a) Graph of an example log prior, $\log p_\theta(\theta)$ (solid horizontal sinusoid curve), as a function of lighting direction θ. Given two hypothetical neurons with preferred lighting directions θ and $\bar{\theta}$, their responses are determined by their log probability density functions, $\log p(\theta)$, $\log p(\bar{\theta})$, indicated by the vertical dashed and solid curves, respectively. For a given stimulus, the larger of the two observed values from the probability density functions $\log p(\theta)$ and $\log p(\bar{\theta})$ determines the lighting direction assumed by the observer, and this, in turn, determines the concave/conxex observer response. These two observed values are noisy estimates of the pdf means, so the choice probability q (see Equation 19) is determined by the relative overlap of the probability density functions (vertical dashed and solid curves) for these two quantities.
b) A log prior with amplitude variations k times smaller than in (a) leads to the same choice probabilities as in (a) provided the noise level σ is also reduced by a factor k. (For simplicity this analysis assumes a concavity preference of 0.5).

any stochastic variation in responses must be a result of noise in the internal encoding of variables used in the decision process, noise which we have modeled by the parameter σ.

Estimating The Discrimination Parameter. Our estimate of the prior depends on the value of the discrimination parameter σ, and we have not addressed how to fix a value for σ. This parameter cannot be estimated directly from experimental data because, for any given value of σ, there is a prior which fits the observed data equally well, as shown in Figure 5. This ambiguity is unavoidable for judgment tasks that depend only on likelihood ratios, which comprise the majority of choice tasks (Green & Swets 1966). For the task considered here, this dependence is made explicit in Equation (17), where the posterior probability is seen to be a function of the ratio L/σ, so that smaller log likelihood differences L can always be reliably detected by using a smaller value for σ.

We have chosen a value $\sigma = 2$dB to analyse human data, which is a generous approximation to the 1dB assumed as a nominal value for the discrimination threshold for human judgments (Jaynes 2003). We anticipate that an analysis similar to that proposed by Jazayeri & Movshon (2006) based on Poisson statistics of individual model neurons would constrain the value of σ.

We note that this choice of discrimination parameter gives estimated lighting priors (see Figure 4) which are similar in shape to the von Mises distributions assumed in (Mamassian & Landy 2001), but which are less localised than implied by the values of their estimated concentration parameters. This is consistent with our aim to use the prior of least commitment, consistent with the behavioural data.

Priors For Other Parameters. The method described for estimating perceptual priors can, in principle, be applied to a variety of other parameters. These include priors for low-level parameters (e.g. speed, direction, line orientation, colour, spectral illuminance), but could also be extended to high-order parameters (e.g. faces, words, syllables).

More Complex Priors. In this study, we have just two variable parameters, light direction (θ) and the convexity/concavity (c) of a fixed shape, and there is no reason to expect these parameters to be correlated in the physical world. Hence we were able to assume independence, and factorize the joint prior $p(\theta, c) = p_\theta(\theta)p(c)$. This assumption was essential in order to make the estimation problem tractable, but it may not be justfied in general.

A prior is just the re-scaled marginal distribution of a multivariate prior distribution. In this study, we have kept all parameters constant except light direction (θ) and the convexity/concavity (c) of a fixed shape. This implies that the prior we have estimated is the marginal of a two-dimensional joint distribution $p(\theta, c)$. Moreover, this joint distribution is itself a marginal distribution of a high-dimensional prior distribution with axes that include parameters such as shape, illuminance spectrum, multiple light sources, colour, and stereo disparity. Had we the time and the means to find the light-from-above marginal of this high-dimensional prior distribution, it is possible that the result would be quite different.

Conclusion

If Helmholtz was correct in stating that perception is a form of "unconscious inference" (Helmholtz 1867) then this implies the existence of a posterior (which determines a perception), a likelihood function (the conditional probability of the retinal image), and a prior (the observer's expectations about the statistical structure of the visual world). Studies in computational neuroscience suggest that the visual system is adapted to the statistical structure of its physical environment (Olshausen & Field 2004). Moreover, this adaptation occurs over a range of time scales, and shapes the evolution of the visual system over generations, and the transfer functions of visual neurons over a matter of seconds (Rieke et al. 1996). Here we have described a method for characterizing the prior for lighting direction. We anticipate that this method will be used to characterize many other priors used for perceptual inference.

Acknowledgments: Thanks to Stephen Isard for useful discussions.

Appendix 1: Experimental Methods

Participants: There were 8 observers, in the age range 21-26 years (mean age = 22.7). Observers all gave their informed consent and were paid £5 sterling.

Apparatus and Procedure: The experiment was run in a dimly lit room. Stimuli were generated using the MatLab (V7.3.0 R2006b) PsychToolbox (V3.0.8) (Pelli 1997). The observer viewed stimuli on a 17" TFT monitor, at a distance of 57cm, using a chin-rest. Each observer completed 576 trials in a morning and afternoon session (not on the same day), making a total of 1152 trials. Stimuli were presented in 16 blocks of 36 trials each. After each block of 36 trials, the observer was able to take a break. In each trial, a stimulus was presented with one of the discs marked with an 'x' in the outermost corner, as in Figure 1. The observer's task was to indicate whether the marked disc appeared to be convex or concave by pressing one of two response keys. Each stimulus remained on the screen until the observer made a response, after which the screen went blank, and there was a pause of 0.5-1s before the next stimulus appeared. Observers received no feedback.

The lighting direction adopted one of 36 directions 'around the clock', at intervals of 10°. For each lighting direction, the stimulus had two complementary configurations. In one configuration, the top left and bottom right discs were convex, whereas the top right and bottom left were concave, and in the complementary configuration it was the other way around. The reason for having two configurations per lighting direction was to ensure that each stimulus looked identical to its complementary configuration when lit from 180° further 'around the clock'. Each disc position (e.g. top left) in each configuration was presented twice at each lighting direction, making a total of 1152 trials (i.e. 4 positions $\times$ 2 configurations $\times$ 2 repeats $\times$ 36 light orientations $\times$ 2 sessions).

Appendix 2: Estimating Lambda

Cross-validation consists of splitting each observer's data into two subsets, a training data set s_{train} and a test data set s_{test}. For each putative value of $\lambda = \lambda_j$, the training data s_{train} was used to estimate $\hat{\mathbf{q}}$ (and therefore the prior) by minimising E. Setting $q_i = \hat{q}_i$ in Equation (23), $E_f(s_{test})$ was then evaluated using the test data s_{test}, which yields estimate of the likelihood of the test data for λ_j. This procedure is repeated over a range of values for λ_j, and the value of λ_j which minimises $E_f(s_{test})$ is taken to be $\hat{\lambda}$. In order to obtain a robust estimate for $\hat{\lambda}$, this whole procedure was repeated using four runs, as follows. Initially, the data were split into four subsets. On each run, three subsets were combined to make the training set, and the

remaining subset was used as the test set. Each of the four subsets took its turn as the test set on exactly one run, with the remaining three subsets being used as the training set. Each run yielded a curve for $E_f(s_{test})$ as a function of λ, and these four curves were averaged. The value of λ corresponding to the minimum of this average curve was taken to be $\hat{\lambda}$ for a single observer. The value of σ was set to $\sigma = 1$ for the simulated observer, and to $\sigma = 2$ for human observers.

Appendix 3: The Mean Vector

The mean vector is the mean of the estimated prior distribution. The direction of this vector indicates the direction of the bias (anisotropy) in the prior, and its length shows the amount of bias. The x and y components of the mean vector are $x = \sum_\theta p_\theta(\theta) \cos\theta$ and $y = \sum_\theta p_\theta(\theta) \sin\theta$.

References

Adams, W., Graf, E., & Ernst, M. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience*, 7(10):1057–1058.

Bishop, C. (1996). *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, UK.

Brewster, D. (1847) *Edinburgh Phil. Trans.* 15, 657.

Green, DM & Swets, JA. (1966). *Signal detection theory and psychophysics*. John Wiley and Sons.

Fisher, N. I. (1995). *Statistical analysis of circular data*. Cambridge University Press, Cambridge.

Gold, J. & Shadlen, M. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cog Sci*, 5:10–16.

Helmholtz, v. H. (1867). *Handbuch der physiologischen Optik*. Voss, Leipzig.

Jaynes, E. (2003). *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge.

Jazayeri, M. & Movshon, J. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, 9:690–696.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transaction of the ASMEJournal of Basic Engineering*, pages 35–45.

Kleffner, D. A. & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, 52(1):18-36.

Körding, K. & Wolpert, M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427:244–247.

Mamassian, P. & Landy, M. S. (2001) Interaction of visual prior constraints. *Vision Research*, 41:2653–2668.

Mamassian, P. & Goutcher, R. (2001) Prior knowledge on the ilumination position. *Cognition*, 81:B1-B9.

Mele, A. & Rawling, P. (2004). *The Oxford Handbook of Rationality*. p149 Oxford University Press US, Oxford.

Olshausen, B. & Field, D. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14:481–487.

Oppel J. J., (1856). Uber ein anaglyptoskop *Poggendorffs Annalen der Physik und Chemie*, 99, 466-469 (Series 11).

Paninski, L. (2006). Nonparametric inference of prior probabilities from Bayes-optimal behavior. In Weiss, Y., Schölkopf, B., and Platt, J., editors, *Advances in Neural Information Processing Systems 18*, pages 1067–1074. MIT Press, Cambridge, MA.

Pelli, D. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10:437–442.

Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature*, 317:314–319.

Rieke, F., Warland, D., van Steveninck, R., & Bialek, W. (1996). *Spikes: Exploring the Neural Code, a Bradford Book*. MIT Press, Cambridge, MA.

Rittenhouse, D. (1786). Explanation of an optical deception. *Transactions of American Philosophical Society*, 9(4):578–585.

Stocker, A. & Simoncelli, E. (2006a). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, 9(4):578–585.

Stocker, A. & Simoncelli, E. (2006b). Sensory adaptation within a Bayesian framework for perception. *Advances in Neural Information Processing Systems Conference*, 18:1291–1298.

Sun, J. & Perona, J. (1998) Where is the Sun? *Nature Neuroscience*, 1(3):183–184.

Weiss, Y., Simoncelli, E., and Adelson, E. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6):598–604.

100 word summary

Perception is unreliable because the retinal image cannot unambiguously specify 3D shape. For example, the picture of a crater is perceived as a hill if the picture is inverted. This dramatic perceptual shift happens because, although visual systems assume that light can come from any direction, some directions are more favoured or probable. The amount of favour assigned to each direction defines a prior probability distribution. We describe a method for obtaining a non-parametric estimated prior, and find that observers favour directions to the left of overhead. This method could be used for priors of speed, colour, or even faces.